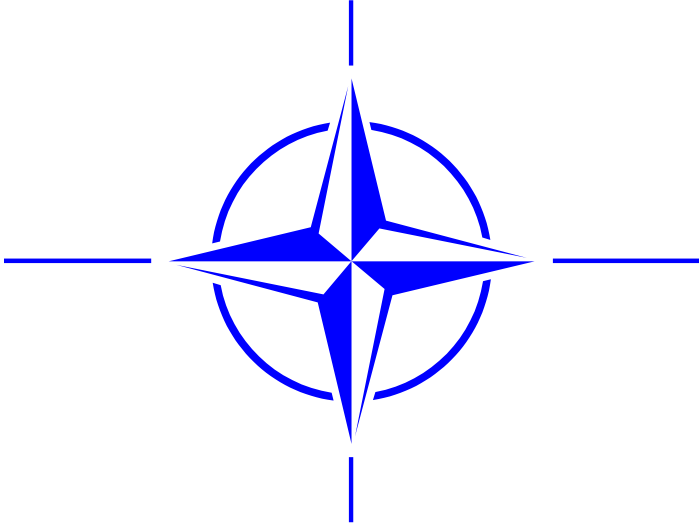


NATO UNCLASSIFIED

VERSION 3.0



NATO OPERATIONS
ASSESSMENT HANDBOOK

01 July 2015

NATO UNCLASSIFIED

Foreword

This handbook defines and describes operations assessment principles, procedures and techniques in NATO, within the construct of operations planning, and should be used in conjunction with the Comprehensive Operations Planning Directive (COPD).

The intended audience for this handbook is staff officers involved with planning, execution or operations assessment in their daily work. Chapter 0 (Introduction), and Annex A (Operations Assessment in NATO) contain general information useful to staff throughout the headquarters. The remaining chapters and annexes are primarily for staff officers and analysts working in operations assessment billets.

The Bi-SC Operations Assessment Sub-Working Group is responsible for the maintenance of this handbook, and is willing to receive lessons identified and comments from operations, exercises or experiments that can be used to validate and improve the guidance offered herein¹.

¹ Comments and questions should be directed to the custodian for the handbook, Andrew Bell at HQ MARCOM, Northwood, UK. (a.bell@mc.nato.int).

Table of Contents

0.	INTRODUCTION	0-1
0.1	Background	0-1
0.2	Hierarchies	0-2
0.3	The Rationale for Operations Assessment.....	0-2
0.3.1	Improving military planning	0-3
0.3.2	Improving decision making	0-4
0.3.3	Structuring and developing knowledge and learning ...	0-4
0.3.4	Supporting strategic communications	0-4
0.3.5	Anticipating and forecasting	0-5
0.3.6	Promoting a comprehensive approach	0-5
0.4	The Operations Assessment Process	0-5
0.5	Definitions and Use of Terms.....	0-7
0.6	The Application of Operations Assessment	0-8
0.7	TOPFAS Campaign Assessment Tool.....	0-11

STAGE 1. ASSESSMENT DESIGN AND SUPPORT TO PLANNING1-1

1.1	The Dependence on Systems Analysis and Knowledge Development	1-1
1.2	Operations Assessment Methodology.....	1-3
1.2.1	Identifying Results	1-4
1.2.2	Monitoring Strategic and Operational Risks	1-4
1.2.3	Monitoring Activity	1-5
1.3	Metric Data Types	1-5
1.4	Metric Data Categories.....	1-7
1.4.1	Factors in Choosing Data Types and Categories	1-8
1.5	Metric Selection and Design.....	1-11
1.6	Measures of Effectiveness (MOE)	1-13
1.6.1	Definition of MOE	1-13
1.6.2	Developing MOE	1-15

1.7	Measures of Performance (MOP).....	1-16
1.7.1	Definition of MOP	1-16
1.8	Application of Metrics	1-18
1.9	Setting Goals.....	1-18
1.9.1	Acceptable Condition (AC)	1-18
1.9.2	Rate of Change (ROC).....	1-19
1.9.3	The Importance of Goals	1-19
1.10	Planning Considerations	1-20
1.11	Causality: A Cautionary Note.....	1-21

STAGE 2. DEVELOPMENT OF A DATA COLLECTION PLAN.....2-1

2.1	Process Overview	2-1
2.2	Data Instruments	2-2
2.3	Data Collection Methods	2-5
2.4	Assigning Responsibility for Data Collection	2-8
2.5	Data Collection Timelines	2-9
2.5.1	Data Collection Frequency	2-9
2.5.2	Data Reporting Frequency	2-9
2.6	Incorporating the Data Collection Plan into the Operational Plan.....	2-10

STAGE 3. DATA COLLECTION AND TREATMENT.....3-1

3.1	Data Collection.....	3-1
3.2	Use of Data Collected.....	3-2
3.3	Duplication of Reporting and Data Exchange	3-3
3.4	Data Treatment	3-4
3.4.1	Procedures for Absence of Data.....	3-5
3.5	Other Applications of the Assessment Data.....	3-6

STAGE 4. ANALYSIS, INTERPRETATION AND RECOMMENDATIONS.....4-1

4.1	Analysis.....	4-1
4.1.1	Data Analysis	4-2
4.1.2	Identification of Change and Trends.....	4-3
4.1.3	Aggregating Analysis.....	4-6
4.1.4	Archiving	4-7
4.2	Reporting and Data/Result Visualisation.....	4-7
4.2.1	The Use of Colour Coding	4-13
4.3	Revalidation of Assumptions, Metrics, and Targets	4-16
4.3.1	Operations Assessment of Effect vs Action.....	4-16
4.3.2	Assessment of Other Planning Elements	4-18
4.3.3	Unintended Effects	4-18
4.3.4	Goals	4-18
4.4	Art of Assessment: The complete assessment	4-18
4.5	Reasons to Reconsider the Operations Assessment Plan.....	4-20

ANNEX A. OPERATIONS ASSESSMENT IN NATO A-1

A.1.	Operations Assessment at the Military Strategic Level. A-1
A.1.1.	Organisation, Roles and Responsibilities at the Military Strategic Level. A-2
A.1.2.	Characteristics of Operations Assessment at the Military Strategic Level
A.1.3.	Summary – Operations Assessment at the Military Strategic Level. A-6
A.2.	Operations Assessment at the Operational and Tactical Level
A.2.1.	Organisation, Roles and Responsibilities at the Operational and Tactical Level..... A-9
A.2.2.	Operations Assessment Process at the Operational and Tactical Level
A.2.3.	Summary – Operations Assessment at the Operational and Tactical Levels
A.3.	Interrelations between Levels of Command
	A-13

A.3.1.	NATO Headquarters and SHAPE.....	A-13
A.3.2.	SHAPE and Operational Headquarters.	A-14
A.3.3.	Operational and Tactical Headquarters	A-14
A.4.	Summary of Operations Assessment at the Military Strategic, Operational and Tactical Level.....	A-15

ANNEX B. CIVILIAN CONSIDERATIONS IN OPERATIONS ASSESSMENT..... B-1

B.1.	Assessment of non-military domains.....	B-1
B.2.	Purpose and Scope of Considering Civilian Assessments.....	B-3
B.3.	Important Factors to Consider	B-4
B.4.	Non-NATO Civilian Actors and the Civilian Sector	B-5
B.5.	Civilian Sector and Operations Assessment	B-8
B.6.	Civilian Approaches to Operations Assessment.....	B-9
B.7.	Comparison of Civilian and Military Approaches.....	B-13
B.7.1.	Origin of Results-Based Management and The use of “Effects” in Operations Planning.....	B-13
B.7.2.	Comparison of Civilian and NATO Terminology	B-16

ANNEX C. OPERATIONS ASSESSMENT – CHALLENGES, KEY MESSAGES AND TOP TIPSC-1

C.1.	The Challenges of Operations Assessment	C-1
C.1.1.	Institutional factors	C-1
C.1.2.	Requirement for comprehensive assessment	C-2
C.1.3.	“Over-metricising”.....	C-3
C.1.4.	Assessment in multi-actor environments	C-3
C.1.5.	Data collection	C-3
C.1.6.	Establishing reliable baselines	C-4
C.1.7.	Methodological challenges	C-4
C.1.8.	Use of assessments by decision makers	C-5
C.2.	Operations Assessment Best Practices.	C-5

C.2.1.	Best Practices Guide for Conducting Operation Assessments for Counter-Insurgencies	C-6
C.2.2.	Key Messages and Top Tips	C-9

ANNEX D. EXAMPLES OF METRICS.....D-1

D.1.	Examples of Metrics	D-1
D.1.1.	Example Metric 1	D-1
D.1.2.	Example Metric 2	D-2
D.1.3.	Example Metric 3	D-4
D.1.4.	Example Metric 4	D-7
D.1.5.	Example Metric 5	D-9

ANNEX E. POLLING TO SUPPORT OPERATIONS ASSESSMENT.....E-1

E.1.	Polling Initiation, Management and Oversight	E-3
E.1.1.	Have You Polled Enough People?	E-4
E.1.2.	Whose Opinions Are You Examining?	E-5
E.1.3.	Difficulties In Accessing Participants	E-5
E.1.4.	Data Collection	E-5
E.1.5.	Questionnaire Construction	E-6
E.2.	Questionnaire Piloting and Testing	E-7
E.2.1.	How Should Participants Be Interviewed?	E-7
E.2.2.	Are the Questions Culturally Sensitive?	E-7
E.3.	Quality Controls.....	E-8
E.4.	Data Analysis.....	E-8
E.5.	Interpreting Results.....	E-8
E.6.	Presenting Results.....	E-9
E.7.	Triangulation.....	E-10
E.8.	Summary	E-10

ANNEX F. GUIDANCE FOR ANNEX OO IN THE OPERATIONAL PLAN..... F-1

F.1.	Introduction.....	F-1
-------------	--------------------------	------------

F.2. Annex OO Template F-1

**ANNEX G. OPERATIONS ASSESSMENT –
REFERENCES.....G-1**

ANNEX H. ACRONYM LIST, GLOSSARYH-1

0. Introduction

Operations Assessment

The activity that enables the measurement of progress and results of operations, and the subsequent development of conclusions and recommendations that support decision making.

0.1 Background

Crises, conflicts and disasters in most cases are characterised by dynamic and unpredictable factors: political transitions, power vacuums, violent insurgencies, terrorism, crime, insecurity, ethnic tensions and conflict, government collapse, economic breakdowns, corruption, disease, and major social upheaval. In response to these factors, intervening military, aid or development actors engage in a wide spectrum of activities to meet often overlapping or conflicting aims. A key question for these actors is how to measure the success of an intervention. In recent decades, an international consensus has emerged that any organisation intervening in a complex crisis needs to have a review and feedback process in order to determine the effectiveness of the intervention and make recommendations for changes—NATO and other military organisations are no exception.

In NATO, this feedback process is called ‘Operations Assessment’ and is critical to inform on progress being made in creating desired effects and achievement of objectives, which in turn allows for adjustments to be made to the plan, and military and political leadership to make informed decisions. Operations assessment both draws from and provides an important input into the systems analysis and knowledge development processes, which build up and maintain a holistic understanding of the situation and operating environment across the Political, Military, Economic, Social, Information and Infrastructure (PMESII) domains.

By necessity, operations assessment will primarily be based on analysis of the elements identified as critical in a military plan and can provide indications of trends in a system’s behaviour, although it can also take a broader look at the operational environment. A final operations assessment product will integrate both the knowledge

derived from this analysis with the Commander and staffs' intuition, experience and judgement.

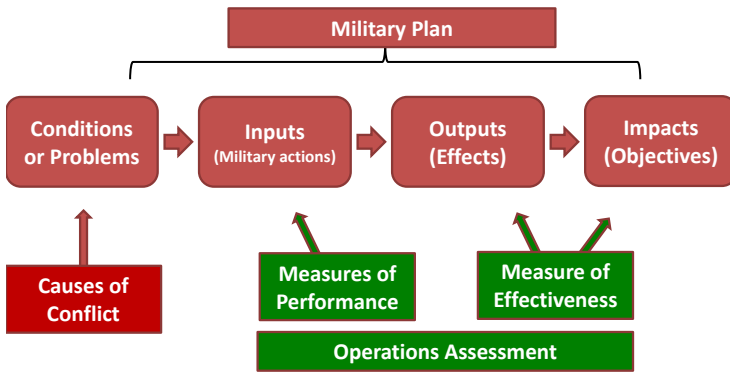


Figure 1 - Operations Assessment in the context of military planning

0.2 Hierarchies

The Command and Control (C2) of Military Operations is conducted using four major interdependent processes: knowledge development, planning, execution, and operations assessment. These processes are described in detail in the Comprehensive Operations Planning Directive (COPD). A reader of this handbook should have a basic knowledge of the planning process used by NATO, and the associated terminology as described in the COPD. Additionally, the COPD should be read in conjunction with this handbook to clearly understand how operations assessment is organised and conducted at the Military Strategic, Operational and Tactical Levels, the interactions between Levels of Command and the interactions with other elements of operations planning and execution.

0.3 The Rationale for Operations Assessment

The idea of measuring progress against defined results has a long history in the corporate, government, and non-profit sectors. The underlying philosophy is that measurement of progress towards results allows informed decisions—“*what can be measured can be*

managed.”² Likewise, in order to make informed decisions Commanders must understand the current situation and the success, or otherwise, of previous efforts in order to plan and direct future activities. Operations assessment is designed to support a continuous cycle of data collection enabling periodic assessments of progress towards desired outcomes, promoting a long-term perspective and making recommendations for adjustments when necessary.

Beyond measuring progress, informing decisions and focusing planning there are several other key reasons for conducting operations assessment, described in the following sections:

0.3.1 Improving military planning

“Military command at all levels is the art of decision-making, motivating and directing to accomplish given missions. It requires a vision of the desired outcome(s), an understanding of concepts, mission priorities and allocation of resources, an ability to assess people and risks, and involves a continual process of re-evaluating the situation.”³ Operations assessment is the continual process of determining if the operation is being executed as planned and if the operation is achieving the desired result(s).

Military actions are guided by plans, which are underpinned by assumptions about cause and effect. The ever-changing environment in which military operations take place means that continuous monitoring and adjustment of the plan is required, either to keep it on track, or to recognise when a situation has changed sufficiently to require plan adjustments. Operations assessment attempts to confirm the validity of planning assumptions and sets a baseline against which change in the situation can be measured and evaluated.

Operations assessment improves the development process of military plans. The act of describing planned objectives in measurable terms gives planners a deeper understanding of system elements affected by activities conducted in military operations, and may reveal interrelationships in the operational environment where metrics overlap. Finally, it prevents “planning by headlines,” in which a

² A popular meme, often incorrectly attributed to Peter Drucker

³ AJP-01(D)

military plan contains sweeping statements about desired changes because it requires planners to describe abstract goals (e.g. “security”) in specific and measurable terms.

0.3.2 Improving decision making

Operations assessment provides information to a Commander to support evidence-based decision making about desired outcome(s), concepts, mission priorities and allocation of resources, people and risks. The assessment process checks whether objectives are met and whether actions have been conducted as planned (timeline and resources). Consequently, operations assessment supports a wide range of decisions: Is the force large enough? Is it in the right place doing the right things? Are new or enhanced capabilities required? Should the plan be modified? Can military strategic objectives be accomplished?

0.3.3 Structuring and developing knowledge and learning

Experience suggests that in the absence of a central enduring operations assessment process, each military rotation tends to develop bespoke assessments rather than evolve the existing one, an approach which can undermine trend analysis. Operations assessment uses structured methods to gather and collate evidence over the duration of the operation. The data collected can help staffs understand the operating environment and how military actions contributed to the success or failure of a mission. It preserves an institutional memory that can be used to learn from each other’s experiences. It supports observations and analysis leading to lessons learned and development of best practice.

0.3.4 Supporting strategic communications

Information operations use public diplomacy and national and international media to influence the perceptions and actions taken by the adversary and other constituencies. Operations assessment can provide credible evidence that can be used to support the Commander’s information operations campaign, and help identify actions to counter adverse media or information campaigns used by the enemy.

0.3.5 Anticipating and forecasting

Operations assessment should extrapolate anticipate developments in the environment and identify the potential impact of these developments on the execution of the plan and the achievement of the desired result(s). This supports planning decisions especially in complex organisations, where decision making takes time. Anticipatory thinking can ensure that resources and capabilities are available when needed.

0.3.6 Promoting a comprehensive approach

Operations assessment is a key enabler in NATO's support to a comprehensive approach for current and future operations. The structured process enables the military to keep track of and share information with other non-military actors. The result is a better understanding of the interconnections and interdependencies between military and non-military activity, and how they should be coordinated and synchronised.

0.4 The Operations Assessment Process

The operations assessment process involves four stages:

- a. Assessment design and support to planning.
- b. Development of a data collection plan.
- c. Data collection and treatment.
- d. Analysis, interpretation and recommendations

Each of these stages will be described in subsequent chapters of this handbook.

As a summary, the basic elements of planning theory are presented in the table below. In planning theory, there are five, and only five elements that constitute a planning typology; however, only the first two are absolutely necessary to define a plan. In other words, while “objectives” and “effects” have different uses in a plan, in conceptual terms they are same (a change in system state or a “result”), and are treated identically by operations assessment.

Planning Elements	Description
RESULTS ⁴	<p>Describe the specific goals in terms of change in the state of a system that you intend to achieve. In military planning terminology, they are:</p> <ul style="list-style-type: none">• End State• Mission• Objective• Decisive Condition• Effect <p>(You do not have direct control over these, and you may only hypothesise that completing certain ACTIVITIES will lead to achieving the planned RESULTS.)</p>
ACTIVITIES	<p>Describe what you intend to actually do, to achieve the stated RESULTS. In military planning terminology, they are:</p> <ul style="list-style-type: none">• Action• Activity• Task <p>(You have direct control over what you actually do.)</p>
FRAMEWORK LINES	<p>Administrative concepts that organise a series of ACTIVITIES and RESULTS by a certain criteria, which may be: Time, type, geographic location, domain etc. In military terminology, these are:</p> <ul style="list-style-type: none">• Line of Operation• Line of Engagement

⁴ Definitions or proposed definitions of each of these terms can be found in the COPD.

Planning Elements	Description
METRICS	<p>The means by which progress towards completion of ACTIVITIES and achievement of RESULTS can be measured. In military terminology, these are:</p> <ul style="list-style-type: none"> • Measures of Performance (MOP, to measure ACTIVITIES) • Measures of Effectiveness (MOE, to measure current system state and thus, looking at change over time, determine RESULTS) <p>A metric measuring a RESULT generally must also include explicit <i>goals</i> (see section 1.9) against which to judge the achievement of results.</p>
SYSTEM ELEMENTS	<p>The functionally, physically or behaviourally related group of regularly interacting or interdependent elements, which forms a unified whole. System elements are specific physical, functional or behavioural entities within a system. Examples of system elements are: people, organisations, facilities, forces, information, processes and conditions.</p>

Table 1 Basic Element of Planning Theory

0.5 Definitions and Use of Terms

Operations Assessment: The activity that enables the measurement of progress and results of operations in a military context, and the subsequent development of conclusions and recommendations that support decision-making.

Metric: A system of measurement, which for the purposes of operations assessment can be an MOE or an MOP as defined by all of its parameters (instrument, Acceptable Condition, Rate of Change, Criteria etc. as described in this handbook).

Measure of Effectiveness (MOE): A metric used to measure a current system state.

Measure of Performance (MOP): A metric used to measure the accomplishment of actions.

Other uses of the term “Assessment”: At this point it is necessary to warn the reader that the word ‘Assessment’ has multiple uses and meanings in NATO. “Assessment” is used in the following contexts that are different from the use considered in this handbook and this will be further explained later in this handbook:

- a. Assessment of the Crisis Situation (NATO Crisis Response System Manual).
- b. Supreme Allied Commander Europe’s (SACEUR’s) Strategic Assessment (Chapter 3 of the COPD).
- c. Uses in Intelligence (Allied Joint Publication 2.0).
- d. Initial Assessment (Chapter 3 of the COPD).

Risk Assessment: The continuous monitoring of strategic and operational risks at the corresponding level of command (Chapter 5 of the COPD).

0.6 The Application of Operations Assessment

Operations Assessment at the Military Strategic Level of Command: At the military strategic level, the term “Operations Assessment” refers to the development and conduct of the measurement of progress and results on the engagement space, although measurements of system state can begin as soon as key elements to monitor are identified, before a plan is written and approved (establishing a baseline). At the military strategic level of command, operations assessment is a function that involves varying combinations of:

- Continual measurement of system states to determine the achievement of strategic effects and progress towards the achievement of strategic objectives in the military domain;
- Continual measurement of system states to determine the achievement of strategic progress and results in non-military domains but resulting from military activity;
- Measurement of system states to determine the achievement of strategic progress and results of activities of non-military organisations;

resulting in:

- An overall evaluation of progress towards the NATO end-state;

- The subsequent development of conclusions and recommendations that support military strategic decision making for the strategic military Commander, and informing the North Atlantic Council (NAC).

Operations Assessment at the Operational Level of Command:

At the operational level, operations assessment focuses on the measurement of two distinct aspects of an operation; the generation of the intended effects (medium-term assessment), and the achievement of the objectives (long-term assessment). These two aspects, although closely related, are the subject of two separate assessment activities.

Medium-Term Assessment seeks to answer the question “*Is the operation being executed as planned?*”. It specifically measures progress in the generation of desired effects towards the achievement of Decisive Conditions (DC) as stipulated in the operational level Operational Plan (OPLAN). The medium-term operations assessment is an integral part of the Commander’s Decision Cycle, and is performed regularly (i.e. once per Battle Rhythm). It consists of a mid-term review of current actions and their desired effects leading to DC along particular Lines of Operation (LOO). It assumes that the approved operational design is correct and therefore measures the progress in generating the planned effects toward the achievement of the planned DC. It uses a combination of MOP and MOE to generate the necessary data and evidence to formulate the assessment. The main output of this assessment are recommendations to the Commander for the adjustment of own actions aimed at generating the effects more efficiently and effectively and indirectly the achievement of the DC more efficiently and effectively. It could also result in recommendations to adjust the effects leading to DCs. The changes resulting from this type of assessment will generally be promulgated through a Fragmentary Order (FRAGO) or a Joint Coordination Order (JCO).

Long-Term Assessment seeks to answer the question “*Is the operation achieving the desired result?*”. It specifically measures progress towards the achievement of the objectives. The long-term assessment is conducted at longer intervals (over several medium-term assessment cycles) to allow enough time for the operational

environment to evolve sufficiently following the creation of planned effects and to draw meaningful conclusions about the adequacy of the plan to achieve the objectives. The long-term operations assessment consists of a long-term review of the achievement of the operational objectives, and seeks to answer the questions:

- If the DC are being achieved but the operational objectives are not, how should the operation be adjusted?
- What influences, if any, outside of the military domain, are limiting or preventing achievement of the operational objectives?
- Are the military strategic objectives and the operational objectives achievable within an acceptable timeframe?

This involves the monitoring of the broader environment across all PMESII domains where they impact on the military operation. It uses metrics specifically related to the operational objectives to generate the necessary data and evidence to formulate the assessment. The main outputs of this assessment are recommendations to the Commander for changes such as the adjustment of the existing operational design (LOO, DC, effects etc.) and requests to the strategic level to influence other factors or actors outside of the military domain which are preventing the effective and efficient achievement of the operational objectives. The changes resulting from this type of assessment are generally promulgated through a JCO or OPLAN revision.

Operations Assessment at the Tactical Level of Command:

At the tactical level operations assessment measures the achievement of planned DCs, actions, tasks or activities using MOPs for each particular component. In addition, the tactical level may measure the achievement of appropriate DCs and creation of operational effects using MOE where they are solely responsible for delivering the operational results.

For each operation, duties and responsibilities may be shared and exchanged between levels, which will be defined in the assessment annexes of plans.

0.7 TOPFAS Campaign Assessment Tool

In NATO, there is one purpose designed software tool to support operations assessment; the Campaign Assessment Tool (CAT). CAT is part of the Tools for Operations Planning Functional Area Services (TOPFAS) tool suite. CAT is designed to support the measurement of progress towards the planned end-state through measures of effectiveness and performance.

CAT allows operations assessment planning, metric data collection and reporting as well as some limited statistical data analysis including causality and trend analyses. Training on TOPFAS CAT is given by your training team or specialist training can be provided by the NATO Communications and Information Agency (NCIA). Further information can be found at <http://topfas.ncia.nato.int> on the NATO classified network.

Stage 1. Assessment Design and Support to Planning

1.1 The Dependence on Systems Analysis and Knowledge Development

The key to operations planning and operations assessment is the ability to understand the operational environment. Systems analysis is a formalised method that supports and enables this understanding.

Systems Analysis is a specialized portion of Knowledge Development that provides the Commander and staff with a comprehensive understanding of the engagement space, using techniques such as the PMESII approach, and how these elements interact as a system of systems. This understanding enables the Commander and staff, in concert with senior representatives of other actors (Other Government Departments (OGD), Non-Governmental Organisations (NGO)) to identify the most effective use of Political, Military, Civil and Economic instruments to achieve desired effects.

A “system” is a functionally, physically or behaviourally related group of regularly interacting or interdependent elements, which forms a unified whole. System elements are specific physical, functional or behavioural entities within a system. Examples of system elements are: people, organisations, facilities, forces, information, processes and conditions. The goal of Systems Analysis is to gain a comprehensive understanding of the capabilities, behaviour and interaction of various systems in the engagement space.

Systems Analysis identifies a network of systems and system elements, their relations and interactions, and their evolution in space and time. This allows a valid contextual determination of objectives, capabilities to achieve those objectives, possible effects, sequences of effects and likely actions in the operations planning process.

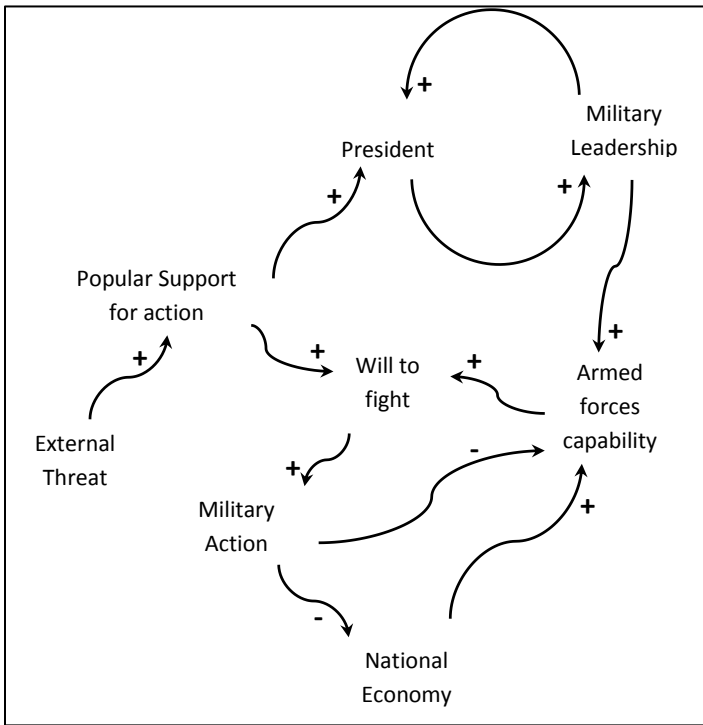


Figure 2 – System Elements and their influence creating a System of Systems

Systems Analysis supports the planning, execution, and operations assessment processes at the strategic, military strategic, operational, and tactical levels, since it provides a comprehensive modelling of the interaction of key actors and systems in the engagement space, although with different levels of granularity. The engagement space is described in varying degrees of complexity, which is directly dependent on the analytical requirements. A tactical view of areas within the engagement space requires more granularity than a strategic, higher level view of the same environment.

An operational plan, in essence, identifies those system elements and relationships between elements that are critical to achieving the desired end state, including those elements or relationships that may not exist at the start of the operation, and determines how to modify, remove or create the necessary elements or relationships such that the

initial system of systems is transformed into the intended end state system of systems.

1.2 Operations Assessment Methodology

Operations assessment measures the state of system elements and, therefore, of system of systems, in order to determine if the desired change is being achieved through the execution of the operation.

The OPLAN will describe the desired changes in state of the various system elements (e.g. external threat, popular support or armed forces capability) and the means by which that is expected to be achieved. The first stage of operations assessment supports the planning process to ensure that the appropriate elements of the plan are measurable. The outcome of the planning process includes an OPLAN operations assessment annex, OO, which describes the means by which one assesses the operation through selecting and monitoring system elements and relationships within the engagement space. Development of the operations assessment annex must begin during the initial planning phases, be refined during the development of the military strategic or operational design and Concept of Operations (CONOPS), and finalised during OPLAN development. As with the rest of the plan, the operations assessment annex will need to be revised and updated throughout the course of an operation.

Within the planning process there is an explicit link between formulating desired future system states (end-state, objectives, effects and DCs) and selecting metrics to measure actual systems states at a particular point in time. There are two types of measurement in operations assessment: measurement of results (change in system state), which uses MOE, and measurement of activity (action accomplishment), which uses MOP. Appropriate metrics may be qualitative or quantitative, subjective or objective, as long as it is possible to define them in sufficient detail that measurements can be made consistently over time.

1.2.1 Identifying Results

Although the planning staff is responsible for defining desired objectives, DCs and effects, they must work in conjunction with the operations assessment staff, who develop the associated MOE. The process of developing MOEs ensures that:

- a. where possible, progress toward desired system states can actually be measured;
- b. the meaning of the system state, as described by the combination of plan element and associated MOE(s), is unambiguous.

This interactive and iterative process may require the modification of currently drafted system states as the understanding of plan and assessment requirements evolve, requiring modification or replacement of effects, DCs or objectives.

Monitoring an MOE over time determines whether results are being achieved according to the plan. If there are elements within the plan that describe outcomes created by other involved non-NATO entities, these elements must be considered as well.

1.2.2 Monitoring Strategic and Operational Risks

Strategic and Operational risks are based on the probability of an operational failure and their consequences. As part of the planning process the Operational Planning Group should review the main factors related to time, space, forces/actors and information within the theatre to identify risky situations and their possible consequences on the accomplishment of an operational military mission. Risks may also be identified resulting from shortfalls in critical capabilities or gaps in coordination with non-NATO actors.⁵

The operations assessment staff may be called upon to monitor important eventualities that pose Strategic or Operational risks and or present opportunities that may require branch or sequel plans. The core aim of operations assessment is essentially to identify risks to the achievement of objectives, establishment of Decisive Conditions or

⁵ Taken from the Comprehensive Operations Planning Directive (COPD) Chapters 3 and 4.

creation of operational effects, i.e. those elements defined in the operational design. Where the operations assessment process identifies differences between the plan and the actual situation within the theatre, these will be highlighted but not necessarily described specifically as risks to the mission.

During the development of Military Response Options the planners will have identified Strategic and Operational Risks, analysed them in a Risk Evaluation Matrix, and determined how the plan can be modified in order to mitigate those risks. Where the level of risk has been determined to be acceptable, either inherently or through the implementation of mitigation measures or it is decided that the cost of mitigating a particular risk outweighs the likelihood or consequences of the risk occurring, there is still a possibility of those risks materialising. As part of the operations assessment process metrics for risks may be formulated and processed as MOE. Monitoring those MOE will then assist in anticipating the likelihood of important risks emerging, and enable proactive decision making when necessary.

1.2.3 Monitoring Activity

As part of the planning process the planning staff will make a number of assumptions about the type and speed of different activities and the size and mix of forces required to conduct them in order to create a desired effect or Decisive Condition (DC). As described in more detail later (see section 4.3) measurement and feedback that allow the comparison of activity undertaken with results achieved is critical in aiding the Commander to make decisions. Operations assessment measures planned activities of importance in the engagement space to allow this comparison to take place, using Measures of Performance (MOP), and monitoring both NATO forces or by working in coordination with other actors (e.g. IOs, NGOs) whose activity is expected to contribute to NATO's desired outcomes.

1.3 Metric Data Types

Metrics require one of four data types. In increasing level of complexity and information content they are:

- **Nominal Data.** Data are organised into categories, where there is no difference in degree or amount between category,

and any ordering by category is arbitrary. Effectively the collection of nominal data is simply a sorting method. For example, friendly forces are categorised by sending nation, e.g. from Albania, Belgium, Bulgaria etc..

- **Ordinal Data.** Data have an order, but has no information about the magnitude of interval between data points. A Likert Scale is a common use of ordinal data, where ‘strongly agree’ represents more agreement than ‘agree’, but without specifying how much more. For example NATO nations ordered by defence spending may have USA 1st, United Kingdom 2nd, France 3rd and so on and so forth, but doesn’t explain how much more the USA spends than UK or France.
- **Interval Data.** Interval data are essentially ordinal data with the extra property that the gaps between the numbers are qualified, or able to be meaningfully added or subtracted. However, an interval scale has no meaningful value for zero, so ratios are meaningless. An example is temperature scales, where 0°C does not mean that there is no temperature. For example, the average daily temperature in Kabul in June may be 25°C, and in December 5°C, so while a difference of 20°C between these months is meaningful, it cannot be stated that June is 5 times as hot as December.
- **Ratio Data.** Ratio data is the most descriptive form of data, where both intervals and ratios are meaningful. Ratio data has a natural zero, unlike interval data, indicating the absence of whatever is being measured. For example the number of personnel in the armed forces of NATO nations is (1999 figures, in thousands) USA 1372, Turkey 639, Germany 322 and so on and so forth. It is valid to say both that Turkey has 317,000 more military personnel than Germany, and that the USA has more than twice as many military personnel as Turkey.

Metrics used in operations assessment are categorised by one of these data types. Knowing the data type is essential in understanding the type of analysis that can be performed, and whether data can be interpreted to draw conclusions, such as, the relative quantity and speed of change in the states of the systems of interest. The ways in which different data types can be manipulated are shown in Table 2.

OK to compute	Nominal	Ordinal	Interval	Ratio
Frequency Distribution	Yes	Yes	Yes	Yes
Median and Percentiles	No	Yes	Yes	Yes
Add or Subtract	No	No	Yes	Yes
Mean, Standard Deviation	No	No	Yes	Yes
Ratio	No	No	No	Yes

Table 2 - Computations with different data types

1.4 Metric Data Categories

In addition to categorisation by type, data can be categorised as quantitative or qualitative, and subjective or objective. There can be misinterpretation of these four terms; therefore the following provides a guide to these terms from an operations assessment perspective.

- **Quantitative:** A number that represents an amount or a count.
- **Qualitative:** An observation that is a word, or a sentence, or a description, or a code that represents a category (attempting to understand rather than prove)
- **Objective:** Facts and the precise measurement of things or concepts that actually exist.
- **Subjective:** Resulting from an individual's personal opinion, experience and judgement.

Metrics used in operations assessment are described by a combination of two of these four categories. All combinations of these two factors are possible, as shown in Figure 3, thus need to be considered in the way the metric is formulated.

	Quantitative	Qualitative
Objective	The number of no-fly zone violations that have occurred in the last week	The mandate to enforce a no-fly zone is approved
Subjective	The air component's assessment of the effectiveness of the no-fly zone, on a scale of 1 to 10	Enemy freedom of action is limited by the no-fly zone

Figure 3. Metric Type Example

1.4.1 Factors in Choosing Data Types and Categories

Quantitative and qualitative can refer to either the underlying data or the means of analysis applied to it. It is possible to apply quantitative analysis to qualitative data and vice versa. Correct appreciation and application is critical in designing valid and reliable assessment frameworks. Objective and subjective are often used alongside quantitative and qualitative with “objective” seen as “good” and “subjective” seen as “bad”. This is overly simplistic.

The styles of quantitative and qualitative analysis are very different. Quantitative analysis is based on measurement of quantity or amount while qualitative analysis is concerned with qualities or types. Quantitative means of analysis are good at answering such questions as who, what, when, where, how many, how much and how often. They provide numerical answers, which can be analysed through statistical methods when appropriate.

Qualitative analysis use different approaches that do not rely on numerical measurements and focus on one or a small number of cases. Qualitative means of analysis are good at answering questions such as why or how and are better at providing an understanding of underlying dynamics and reasons; they can consider why people may think or behave in a certain way while quantitative approaches simply tell you that they do.

Qualitative raw data are usually textual although they can be grouped and aggregated to provide numerical data. The fact that data is expressed in numbers or text does not define it as either quantitative or qualitative; it is important to look behind the expression to see

what is actually being measured. Qualitative data can only be properly analysed with quantitative methods if codified in numbers and there is a large enough sample of cases to allow the use of these methods.

Objective refers to observable facts, things that exist independently of perception and that are unaffected by personal perspectives. Subjective brings in perspectives, attitudes and emotions; it refers to how people think about an issue. Ultimately all assessments are subjective as they involve making judgments and the level of subjectivity increases the further an assessment progresses. Deciding what data to collect and what to omit involves judgment and therefore an element of subjectivity exists in the very design of any assessment framework. The meaning that individuals place on any data also draws on their worldview and previous experiences, often unconsciously and therefore subjectivity increases.

Pure “hard” objective data unaffected by perceptions, institutional norms or attitudes is rare. Assessors therefore need to develop an understanding of the strengths and weaknesses inherent in any data, ensure that conclusions drawn reflect this and the end users of the assessments are aware of them; overly simplistic characterisations of data as being objective or subjective can be misleading and unhelpful.

The optimum data type and category to use to understand the state of a system should be derived from the system of system analysis, and the factors that will influence the relationships between various system elements. For example, data about a personal relationship between two leaders (gathered by people who know them) is likely to be nominal or ordinal, and qualitative and subjective. If a relationship or ability to influence is dependent on a financial relationship or the size of capability of an armed force, then it is most likely that interval or ratio data that is objective and quantitative will be needed.

The terms 'quantitative' and 'qualitative' refer to kinds of data. The definitions of these terms are uncontroversial and can be found in any standard statistics text book. Witte & Witte (2009)¹, for example, presents the distinction concisely, defining quantitative data as follows:

"When, among a set of observations, any single observation is a number that represents an amount or a count, then the data are quantitative."

So body weights reported by a group of students, or a collection of IQ scores, or a list of task durations in seconds, or Likert scale category responses, or magnitude rating scale responses, are quantitative data. Counts are also quantitative, so data showing size of family, or how many computers you own, are quantitative.

Witte & Witte define qualitative data as follows:

"When, among a set of observations, any single observation is a word, or a sentence, or a description, or a code that represents a category then the data are qualitative."

So 'yes-no' responses, people's ethnic backgrounds, or religious affiliations, or attitudes towards the death penalty, the presidential candidate you wish to vote for, or descriptions of events, speculations and stories, are all examples of qualitative data.

Note that the fail-safe way to distinguish between quantitative and qualitative data is to focus on the status of a single observation, or datum, rather than on an entire set of observations or data. When viewed as a whole, qualitative data can sometimes bear a striking resemblance to quantitative data. 57 'yes' responses vs. 43 'no' responses look like quantitative data, but they are not. Although these numbers are important (and essential for some statistical procedures) they do not transform the underlying qualitative data into quantitative data.

<http://www.userfocus.co.uk/articles/datathink.html>

¹Witte, R. S. & Witte, J. S. (2009) Statistics. John Wiley & Sons.

The strengths of quantitative data are that they can be precisely analysed in a descriptive or statistical way. Data can easily be aggregated or disaggregated to provide a broad view of a population or of a small subsection of it and to enable comparisons to be made. Care must be exercised when using quantitative data because of the risk that a mistaken reliance on an incomplete set of metrics may limit the depth of understanding of complex situations, and because subjectivity can be hidden behind seemingly objective data. Furthermore, results of statistical data analysis are still exposed to subjective interpretation with a potential negative impact on the accuracy and reliability of the conclusions drawn.

Qualitative data and means of analysis are good at providing insights into attitudes, beliefs, motives and behaviours; however, care must be taken in drawing conclusions that are considered to be representative of a wider population. It is also more likely to be influenced by factors such as the setting in which the data is collected and the perspective of those conducting the data collection.

Neither quantitative or qualitative, subjective or objective data or means of analysis are better or worse; they all have strengths and weaknesses and good assessment frameworks will draw on a mix.

1.5 Metric Selection and Design

In selecting appropriate metrics, the use of objective and quantitative data will give reliable, high confidence evidence from which to derive assessment products. When a qualitative metric is used categorisation as ordinal data is preferred, and the selection and specification of a range of levels against which the metric can be assessed should be part of the planning process. If it is only possible to use a metric that can be classified as nominal data, then it is effectively the equivalent of a yes/no answer.

When using qualitative data it is necessary to identify intervals (described as levels) that fulfil the criteria of:

- A small enough number of levels that a change is meaningful
- A large enough number of levels that change can be identified in a timescale meaningful to the operation

When using quantitative data it is possible to identify interval or ratio values for the metric, and reference points such as the current situation, desired situation, and comparison to, for example, regional norms to support the assessment process.

Some metrics can be unbound, i.e. it is possible that they can have no theoretical upper or lower limit in value, e.g. in a country affected by a crisis, the number of potential refugees is limited to the size of the population and is therefore bounded, but the number of days for which refugees are displaced from their homes could theoretically be unlimited, and is therefore unbounded. Care should be taken in defining levels for a qualitative metric that is unbounded, as once the highest (or lowest if appropriate) level is reached, it will no longer be possible to identify change in the system of interest, without rescaling the metric and thus undermining all previous analysis and assessment.

As an example, for an effect stating: *“Violent anti-government demonstrations (which involve attacks against security forces) are reduced in province Alpha, by 80% compared with 2010”*, an appropriate MOE is the number of violent demonstrations. It is vital, however, to clarify the interpretation of these statements, noting the following:

- **Quantitative Metric with Ratio data.** The exact number of demonstrations can be recorded together with meaningful goals.
- **Qualitative Metric with Ordinal data.** The number of demonstrations could be described in levels of ‘none’, ‘low’, ‘medium’, or ‘high’. In order to give meaning to the levels a better definition might be ‘none’, ‘manageable by national security forces in normal operations’, ‘manageable by national security forces with enhanced manning levels’ (e.g. leave cancelled), ‘manageable by national security forces only with external support’.

In defining metrics it should always be possible to indicate if the situation deteriorates from the starting point before it improves. For example, taking the violent demonstration example above, if the situation at G day is that external support is required to manage the situation, further levels such as ‘manageable by national security forces with levels of support that detracts external forces from their primary mission’, or even ‘unmanageable even with available external support’, may be appropriate.

In all cases, quantitative or qualitative, detail must be given of the exact interpretation of data within the definition of the metric. For example the exact meaning of ‘violent demonstrations’ should be defined, i.e. 1 incident resulting in arrest or more than 10 incidents, 1 stone thrown or more than 20, ‘violent’ indicated by incidents resulting in casualties requiring medical attention. This again ensures consistency of data over time and changes of personnel.

It is likely that in the decomposition of the systems of interest in the operational environment, higher level elements of the plan such as objectives will cover broad areas of the operation, and thus may require multiple metrics to determine their status, whilst lower level elements such as effects may be specific enough that a single metric is all that is required. This may result in a plan element being described by a combination of quantitative and qualitative metrics; the means for combining these into a single assessment is described in section 4.4. Similarly at higher levels of the command chain the plan may reflect more abstract concepts such as “political will” and “public perception,” and thus more qualitative analysis may be likely than quantitative in order to produce an assessment. It is possible that, for example, a qualitative input to a strategic assessment could be based on the synthesis of a number of quantitative metrics generated by the operational level. The relationship between assessments at the different levels of command is discussed in annex A.

1.6 Measures of Effectiveness (MOE)

Measure of Effectiveness (MOE): A metric used to measure a current system state.

1.6.1 Definition of MOE

An MOE is defined as a metric used to measure a current system state. The MOE helps answer the question “*Are we on track to achieve the intended new system state within the planned timescale?*” Each MOE describes a snapshot of the state of a system element at one particular moment in time; therefore to fully understand how a system is changing this will almost certainly require multiple MOE

for each system, and MOE must be repeatedly measured across time to determine changes in system states. It is the trends that result from these repeated measurements that allow the determination of progress (or lack of) in an operation.

An MOE must:

- a. Describe one system element or relationship of interest.
- b. Be observable and defined in sufficient detail so that assessments are produced consistently over time.
- c. Be derived from data that is collectable within the means and resources of the operation (see Chapter 3).
- d. Describe how the element is expected to change.
- e. Be as specific as possible (ensure you are assessing only and exactly what you want).
- f. Be sensitive to change in a period of time meaningful to the operation.
- g. Be culturally and locally relevant.
- h. Have an associated Acceptable Condition (see 1.9.1).

An MOE should ideally:

- i. Be reducible to a quantity (as a number, percentage, level for ordinal data etc.).
- j. Be objective.
- k. Be cost-effective and not burdensome to the data collectors.
- l. Have an associated Rate of Change (see 1.9.2).

An MOE could possibly:

- m. Be qualitative, describing key elements of a situation as a narrative or statement describing the status of a system element.
- n. Be subjective, noting levels of confidence and validity associated with the source.

1.6.2 Developing MOE

The MOE should be written in concert between the planning and operations assessment staffs during the development of Effects, DCs and Objectives. Based on the results of MOE development, the statements that define these plan elements may require revision. Collection methods must be a key consideration during MOE development to ensure that the desired metric can be measured.

Some considerations for the operations assessment staff during MOE development include:

- While supporting planners in drafting End State, Objectives DCs or Effects, ensure that they can be measured, and their description is written in a manner that can be measured. Participate in the Joint Operations Planning Group (JOPG)/Operations Planning Group (OPG) to ensure changes in system state that are defined in the plan are accompanied by appropriate, workable MOE.
- Consider data instruments for proposed MOE – even if the element can be measured, failure to collect the required data will make it impossible to assess the creation of a desired result. Whenever feasible, plan to use multiple independent data sources to guarantee availability of data and to improve the reliability of the assessment.
- Selection of MOE will require significant input from Knowledge Development (KD) or related systems analysis functions. This input provides deeper insight to ensure that the chosen MOE is actually related to the system element in question.
- The relevance and importance of individual MOE will vary with the phase of the operation and should both respond to, and inform Commanders' priorities and decision making.
- Independent measurement of progress toward the Effects, DCs, Objectives and End State is important. To avoid the trap of assuming causality (see section 1.11) this will require each level in the operational design hierarchy to be measured and assessed independently.

Undesired effects may be identified during the planning process. Undesired effects are those that disrupt or jeopardise the achievement

of objectives; it includes possible negative or detrimental consequences of own-force actions identified in the plan. If plan modifications cannot avoid these Undesired Effects, they should be incorporated in the plan by defining the opposite of the undesired consequence and re-writing them as desired Effects and developing appropriate MOE.

For example: consider an undesired effect identified as “Due to the perceived hostile presence of own force, local militia activity increases.” This may be included in the plan as “Local militia activity remains low” or “Own force not perceived as hostile.”

There may be Undesired Effects that may not be easily transformed into desired Effects, or where doing so may disrupt the construct of the plan. In this case, MOE must still be scripted for the Undesired Effects. This ensures data collection requirements are identified and the system is monitored for undesired changes.

1.7 Measures of Performance (MOP)

Measure of Performance (MOP): A metric used to determine the accomplishment of actions.

Once the hierarchy of Objectives, DCs, Effects and MOE have been approved by the Commander, the planning staff begin development of the Actions necessary to achieve those system states. The operations assessment staff remains involved, developing the required MOP. However, the key consideration here is ensuring that the MOP is directly related to the DC or Action – not to the other elements of the plan.

1.7.1 Definition of MOP

The MOP allows the measurement of activity, intending to answer “*Are the actions being executed as planned?*” If, during execution, progress towards the achievement of desired effects is not made as expected, one possibility is that actions are not being carried out as planned. An MOP is defined as a “metric used to measure the completion of activity, either routine repeating activity or the completion of a specific one-time activity”. Each level, from military

strategic downwards, will normally develop MOP for the actions they will execute or appropriate DCs they will achieve.

An MOP must:

- a. Align to one or more (own-force) actions or DCs.
- b. Be observable and defined in sufficient detail so that assessments are produced consistently over time.
- c. Be derived from data that is collectable within the means and resources of the operation (see Chapter 3).
- d. Describe specifically how the action is expected to be executed or DC achieved.
- e. Be as specific as possible (ensure you are measuring only and exactly what you want).
- f. Be sensitive to change in a period of time meaningful to the operation.
- g. Have a known deterministic relationship to the action.
- h. Have an associated Acceptable Condition (see 1.9).

Additionally, an MOP should ideally:

- i. Be reducible to a quantity (as a number, percentage, level for ordinal data etc.).
- j. Be objective.
- k. Be cost-effective and not burdensome to the data collector.
- l. Have an associated Rate of Change (see 1.9).

An MOP could possibly:

- m. Be qualitative, describing key elements of an action or DC as a narrative or statement.
- n. Be subjective, noting levels of confidence and validity associated with the source.

Examples of MOP are presented in annex D.

1.8 Application of Metrics

It is important to note the key difference between MOE and MOP: The MOP measures the *status* of own-force actions, but does not measure the system state changes of other actors that may *result* from those actions. Changes to the system that occur as a result of own force actions (and other external influences) are measured by MOE. In essence, you have direct control over items measured by the MOP, but no direct control over items measured by an MOE. In other words, MOP measure the amount of effort being *put into* an action, while MOE measure the *outcome or impact* by looking for the changes that result.

Once drafted by the operations assessment staff and agreed by the Commander, the MOE and MOP are integrated into the operations assessment annex (Annex OO) of the OPLAN (See annex F). Any change or update to operations assessment requirements must be implemented through established procedures within the Headquarters (e.g. FRAGO, JCO or OPLAN review as appropriate).

1.9 Setting Goals

The setting of explicit goals for each metric to judge the achievement of results is done through the use of two mechanisms, which should be defined for each metric:

- Acceptable Condition
- Rate of Change

1.9.1 Acceptable Condition (AC)

Acceptable Condition: A defined level for the metric at which a desirable situation has been achieved.

An AC is defined during the planning process as either

- the intended value of a metric, representing a specific element of a system state when an effect is created or objective achieved for MOE.

- the level of activity required in order to make progress at the intended rate during an operation for an MOP.

1.9.2 Rate of Change (ROC)

Rate of Change: A rate of change is the amount of change in a metric over a specific time during an operation

A ROC may be applied in three separate circumstances:

- **Desired (DROC):** A ROC developed during the planning that indicates the intentions of the OPG in creating an effect or achieving an objective in a specified period of time, or in application of resources as they become available in order to undertake an action. There is likely to be a direct relationship between the planner's expectation of the amount of resources devoted to a specific action (Troop-to-Task analysis) and the ROC of the associated MOP for the defined actions and DC and MOE for the resulting effects, DCs and objectives.
- **Actual (AROC):** The ROC measured by the assessment process during the execution of an operation.
- **Predicted (PROC):** The anticipated ROC, which is a useful estimate based on previous history, current intentions and a number of assumptions specific to the circumstances involved.

It is the comparison between the desired, actual and predicted ROCs that will reveal which aspects of an operation are on track, and which require attention.

A ROC does not have to be constant over time; it can change in order to reflect the arrival and availability of appropriate forces, changes in main effort associated with different phases of an operation, or external factors such as the impact of seasonal weather on the ability to conduct operations.

1.9.3 Establishing Goals

Goals may be set by:

- Using reference points such as the restoration of pre-crisis conditions;
- Comparison to other nations in the region or other appropriate groupings;
- Analysis methods such as modelling and simulation of the desired change in system state;
- Reference to historical information from other operations.

The responsibility for the setting of goals (i.e. numerical values and time aspects), remains with the planning staff.

As ROCs and therefore the timeline for the achievement of ACs will in most cases be dependent on the amount of resources devoted to a particular task, the process of selecting the correct combination of AC and ROC will help planners to identify interdependencies and sequencing between different elements of the operations. This in turn will help in setting priorities, and promote the identification of undesirable effects if a system state is not changed in an appropriate timescale (for example, failing to set security conditions that allow sufficient humanitarian aid to be delivered before winter weather arrives will result in an undesirable humanitarian effect). Therefore once a combination of AC and ROC are set this will then directly feed the Troops to Task analysis, which in turn will define the size and scope of the Combined Joint Statement of Requirements (CJSOR) for the operation in an objective and justifiable way.

During the execution phase, analysis may demonstrate that the initial AC and ROC are not appropriate; changes in priorities or phases, improved knowledge of the operational environment, and responses to the changing dynamics of the operational environment are some of the reasons why different values more suitable. As changes to AC and ROC will have an impact on application of resources and campaign synchronisation for example, they should be prepared in conjunction with the planning staff and implemented after command approval.

1.10 Planning Considerations

It is a planning responsibility to define all goals (i.e. Acceptable Conditions, Rates of Change) based on system states; however, the operations assessment staff should support this effort with analytical expertise from Operational Analysts.

Each MOE must include goals (i.e. Acceptable Condition and/or Rate of Change) described by appropriate and meaningful criteria to ensure common understanding of the change in a system element or relationship that indicates its status.

1.11 Causality: A Cautionary Note

Operations assessment is about measuring execution of implemented military actions and the effectiveness – or results – of those actions. By carefully designing metrics to allow activity (MOP) and results (MOE) to be measured, and then collecting data, operations assessment staff will compare the completion of actions with the level of achievement of results.

It may be tempting or seem appropriate to assume that when all associated actions are complete, the effect must be created; or when all effects are created, the objective is achieved; or when all objectives are achieved, the end-state must therefore be reached. Completion of all assigned actions may not lead to creation of the desired effect for many reasons: unknown or unaccounted for actors in the theatre; an unknown linkage with a different system causing an adverse (unwanted) impact; or perhaps not all required actions were identified in the original plan.

In general, avoid the temptation to assume causality. Rather than trying to identify and demonstrate how changes in the environment can be “attributed” to particular actions (implying causal relations), it may be more constructive to talk about how activities might or might not have *contributed* to the creation of effects or objectives.

The ideal solution to this is to independently measure each level in the decomposition of the operational design into specific plan elements; i.e. in addition to assessing the achievement of each effect, the state of the objective that those effects are planned to achieve should also be assessed by independent metrics. Where this is not clearly possible, the use of terms such as “correlation” and “contribution” are much more in line with the realities of what can be accomplished by planning and assessment staffs.

Current thinking in academia on statistical theory and assessment of complex programs is of the view that causality is extremely challenging to infer, in all but the simplest of cases⁶

⁶ See, for example, Sobel, M. E. (2000), Causal Inference in the Social Sciences. *Journal of the American Statistical Association*, 95(450), 647-651. Posovac, E., & Carey, R. (2007). *Program Evaluation: Methods and cases* (7th ed.).

Stage 2. Development of a Data Collection Plan

2.1 Process Overview

After the metrics for the plan have been established, the operations assessment staff is responsible for the development of a data collection plan in co-ordination with the planning staff. This process should be synchronised with other members of staff who will become responsible for collecting data, and who will therefore be members of the Assessment Working Group (AWG – see annex A). Where possible the data collection should exploit existing reporting or data streams. A data collection plan should be agreed that indicates each instrument to be used for MOE or MOP:

- The data parameters, such as;
 - 1) Units of measurement
 - 2) Scale, if appropriate
 - 3) Categorisation for nominal or interval data
 - 4) Upper and lower bounds
 - 5) Specific criteria such as context
- The source of data;
- The method of collection;
- The party responsible for its collection;
- The format in which it should be recorded;
- The required frequency of recording (including key timings);
- The frequency of reporting;
- For processed or interpreted data, whether the raw data should also be supplied (e.g. for archiving)
- A short description of how the data contributes to the assessment process;
- Any other necessary information.

The creation of the data collection plan assists in clarifying the ‘measurability’ of the selected MOE and MOP and may invigorate further revision of the metrics. The data collection plan should always be synchronised and de-conflicted with all reports and returns across the headquarters (HQ). There is an inevitable tension between the desire for the collection of ideal metrics with the practicalities of

what can, and cannot, be measured. It is possible that some metrics are identified as either un-measurable (data required is unreliable or non-existent), or that the effort required to collect the data outweighs the benefit of measuring. The aim should be to provide the most practical mix of metrics possible that achieves sufficient rigour, whilst remaining achievable.

The majority of MOP data will probably be organic—generated, captured, and reported by units within the command structure—while some might be reported by external non-military organisations. In general, collection of data for MOP should commence when the Action(s) start, and stop after the Action is assessed as complete.

Collection of data for MOE is more situation dependant. In some circumstances reporting of progress towards effects not yet scheduled may yield erroneous results. In other circumstances it may be appropriate to collect data for MOE in order to establish a baseline. This should be a continuous process to monitor changes in the system prior to execution. The compilation of data establishes a baseline assessment, which is the capture of current system state just prior to any attempt by own forces to modify the system. By definition this includes assessment of effects prior to execution of any Actions.

The data collection plan specified by the operations assessment staff is published with the final OPLAN or Operations Order (OPORD) in Annex OO (See annex F). In the case where the resources required to collect the data are significant, the planning staff should create separate actions and MOP that reflect this task. Care should be taken regarding the resource allocation cost/benefit required for data collection. Once the OPLAN is approved by the Commander, and prior to commencement of the operation, all levels of Command should start the operations assessment execution process. In order to have the most comprehensive baseline assessment available, and indeed to inform the planning process, some data collection and assessment activity may need to occur prior to OPLAN approval.

2.2 Data Instruments

The operations assessment and planning staff, with input from KD, subordinate Liaison Officers and other actors (including non-military organisations), should specifically identify the expected instruments that will be used to obtain the data. At the Military Strategic level, and possibly at the Operational level, a significant proportion of MOE

data may come from non-military organisations, with implications for relationship building, mechanisms for information exchange, and of course security classification. For example, data could originate from sources such as:

- Local population (formal or informal surveys);
- Host Nation officials (formal or informal surveys);
- Host Nation records;
- Other Government Departments (i.e. Embassies, Development Departments);
- International organisations working in the area (e.g. United Nations, World Bank, International Monetary Fund, European Union, Organisation for Economic Cooperation and Development);
- Non-Governmental Organisations / Private Voluntary Organisations working in area;
- Own force informal and formal observations (e.g. patrols, intelligence);
- Media and other open sources (local, national and home radio, Internet, Social Media, TV and other print sources);
- Commercial Data Sources;
- Other national security forces;
- Legacy records (Lessons learned and historical records of all the above);
- Subordinate assessments;
- Subject Matter Experts (SMEs).

DataCards is an example of an online resource that indexes sources of data that relate to irregular warfare and socio-cultural topics in support of operations assessment and analysis. DataCards initially focused on Afghanistan but has expanded worldwide since then. It provides a summary description and evaluation of the content, quality, intended purposes and potential users of each data source. <https://www.datacards.org>.

Each data source identified requires appropriate scrutiny prior to and during use. Validity of inputs from sources outside Planning and Assessment staffs control must be well understood and reported by the staff. Also, source information should always be linked to the data collected to help provide full data disclosure when reporting. Without this information the credibility of any final assessment could be disputed if the conclusions of the assessment appear to be overly positive or negative when compared with general perceptions and expectations.

It should be remembered that operations assessment is not an end in itself; the data and analysis that it produces must be useable to feed decision making processes across a headquarters. Some data sources such as subordinate commands or external organisations may be using the data they supply in different ways to NATO for their own assessment purposes. Consideration should be given as to whether data being sourced from other organisations is raw or processed, and in the case of processed data then knowledge about the raw data, assumptions and processing methods involved should be obtained. Other organisations may use NATO supplied data in a similar way, and consideration should be given to ensuring that assessments using the same data instruments for different purposes produce assessments which are consistent or mutually supporting where appropriate.

It is important for the operations assessment staff to specify explicitly the expected data instrument including source for the MOE or MOP and identify several back-up or corroborating sources, for the following reasons:

- **Multi-source data is more easily verified** - A data item from one source is not as valuable as when the same data item is corroborated by other sources.
- **Data is subject to human bias** - If the data item involves visual observations (e.g. number of open shops) or perceptive observations (e.g. sense of security in the town), the data may vary significantly depending on the source chosen.
- **Tracking metadata is as important as the data itself** - Regardless of whether data is taken from one or multiple sources, the origin is important for analysis purposes to track the data by source, over time. Where data is supplied by

multiple sources the relative validity of each source should also be reported.

- **Data Archiving through historical records and data backups is essential** – In addition to capturing lessons learned, for long operational deployments, analysts can work to improve metrics by performing trend analysis of data over time. Improved historical data capture can improve the ability to use predictive analytical techniques where opportunities arise. Where it is necessary to change metrics in response to changes in the operational plan or operational environment then the ability to access historical raw data and reanalyse it in a different way will be invaluable in baseline and trend analysis.

2.3 Data Collection Methods

The operations assessment staff should have considered data collection methods during the planning process, and will confirm them whilst populating the data collection plan. The planning staff in collaboration with the operations assessment staff identify resources required to achieve data collection, prepare data collection orders to be given to subordinates, and liaise with civil-military interaction staff (J9, Civil Military Interaction (CMI), or Civil Military Cooperation (CIMIC) for example) in order to identify appropriate liaison with non-military actors to set up data exchange procedures (subject to the appropriate approvals within NATO). There may be cases where an action must be created for the sole purpose of data collection. Inevitably there will be a balance required between the resources allocated for data collection and resources for other military actions.

Table 3 gives some examples of data collection methods and associated advantages and disadvantages.

Data Collection Method	Description	Advantages	Disadvantages
Military Survey	A selective and planned questioning by military forces	<ul style="list-style-type: none"> - Ease of tasking military forces - Superior mobility - Ability to access difficult environments - Good for gathering raw qualitative data 	<ul style="list-style-type: none"> - Response of subjects may be biased by negative perception of military - Military forces may not have specific skills in surveying
Survey	A selective and planned questioning of subjects by non-military parties (e.g. charities, NGOs, or specialists)	<ul style="list-style-type: none"> - Survey by independent bodies can be more impartial - Reduced burden on military - May use survey specialists 	<ul style="list-style-type: none"> - Difficult to task non-military orgs and extra financial cost may be involved - Non-military orgs may have reduced ability to access difficult environments
Focus Group	A group of people are asked about their perceptions, opinions, beliefs towards a subject of interest. Questions are asked in an interactive group setting where participants are free to talk with other participants.	<ul style="list-style-type: none"> - Enables collection of in depth attitudes, beliefs and anecdotal data - Group dynamics facilitate idea generation. - Participants not required to read or write, relies on oral communication. 	<ul style="list-style-type: none"> - Requires strong, trained facilitator - Difficult to make conclusions that represent a population view - Dependant on how representative the group is - Limited ability to repeat data collection over time

Data Collection Method	Description	Advantages	Disadvantages
Structured Interview	A planned, targeted discussion with a subject where the objectives of the discussion are pre-determined and noted in the data collection plan.	<ul style="list-style-type: none"> - Often captures richer information than a survey - Many methods may be used: face to face, email, telephone, video conference, etc. - Provides opportunity to probe and explore ideas in depth 	<ul style="list-style-type: none"> - Interviews are very time consuming - May require trained interviewer/transcriber - Dependant on how representative the individual is
Structured Observation/ De-Brief	A set of specific observations collected during routine work, followed by formal question, or asked to report observations at a specific time.	<ul style="list-style-type: none"> - Good approach to discover behaviours 	<ul style="list-style-type: none"> - Structured observations can be time consuming. - Human bias of difference in perceptions
Military SITREP	A formatted report intended to convey a pre-defined set of information in relation to a specific event or activity, or a routine (time dependant) report.	<ul style="list-style-type: none"> - A standardised set of information which helps in consistency of reporting - A normal part of military business rather than an extra burden 	<ul style="list-style-type: none"> - Limited opportunities for reporting non-standard data, or for changing report formats for mission specific data
Automatic Media Collection	An automated collection and analysis of open source media (e.g. RSS feeds, online market data, social media, country watch reports) or closed source media (NATO / national intelligence)	<ul style="list-style-type: none"> - Resources and time are saved in the efforts required for the data collection 	<ul style="list-style-type: none"> - More analyst time is required to sort through data - Automatic methods may either collect too much or too little, or miss vital data.

Data Collection Method	Description	Advantages	Disadvantages
Manual Media Collection	Manual observation (e.g. reading documents, logging events, photocopying of open source media or closed source media	- The data collection can be more thorough, with a certain amount of analysis being done simultaneously.	- Manual data collection is time consuming

Table 3 Data Collection Methods

2.4 Assigning Responsibility for Data Collection

The operations assessment and planning staff should assign individual persons, units or organisations with responsibility for each data collection item in the data collection plan. This is important for the following reasons:

- Assigning responsibility increases the likelihood of the task being accomplished;
- A reporting chain is clearly identified and communicated;
- In the event of a data query, the analyst can direct questions to the person or organisation responsible for that data item;
- As the source becomes familiar with the data and a deeper understanding is developed more, rich, long-term analysis of the data is possible;
- If data collection actions fall to persons or organisations outside the military domain, action can be initiated to establish links with those particular persons or organisations;
- If data is to originate from an organisation not likely to be compliant, then action to seek an alternative source of data or even an alternative metric can begin.

2.5 Data Collection Timelines

The operations assessment staff, with guidance from the Commander and execution staff, determines the frequencies⁷ for data collection and for reporting⁸. There are two important frequencies to be considered in the data collection plan:

2.5.1 Data Collection Frequency

The number of times per day, week, month etc. that the data should be recorded (though not necessarily observed). For example, the number of attacks on NATO forces in province X might be recorded each day, but the data collection plan might specify ‘Attacks per week’. Some important considerations are:

- Matching the collection frequency with the likelihood of observable change and the decision making cycle that it is feeding. It is likely to be inefficient to record the number of events per hour, when only a few occur each month, or events per day to feed an annual Periodic Mission Review.
- Conversely, if the incidence of a highly contagious disease in a refugee camp was being monitored, daily figures would be more appropriate than the number of new infections each month.

The requested frequency will need to support the most rapid assessment need. For example, assessment of an Action may need to occur daily at the tactical or in-field level. The same raw data may be used at a higher level in the command chain, where a monthly assessment is more appropriate. Specification of the data collection requirements will need to consider these factors.

2.5.2 Data Reporting Frequency

The number of times per day, week, month etc. that the data should be reported back to the operations assessment staff. Depending on the Commander’s requirements, this may vary greatly depending on

⁷ ‘Frequency’ means the number of instances occurring in a certain time period.

⁸ In this sense, ‘reporting’ implies the person or organisation responsible for data collection giving the required raw data at the required time to the designated point of contact (in the assessment staff), without necessarily any analysis on the data.

the nature of the situation. For routine or non-priority items, especially if derived from existing reporting or data streams, the execution staff can indicate an initial frequency, which for these items it may be altered later if initially found unsuitable to support the assessment process.

The data collection frequency and data reporting frequency may not be the same; for example a higher command may want the data to specify the number of ‘attacks per day’, but may only need that reported once per week.

2.6 Incorporating the Data Collection Plan into the Operational Plan

The final data collection plan must be formed into orders published as an annex to the OPLAN, Annex OO – Operations Assessment. This ensures that all levels of the Command know exactly what actions and resources they must commit to data collection. Consideration must be given, however, to the mechanisms required to modify the information published in the OPLAN as the operation progresses and assessment requirements change, for example through FRAGOs. Regardless of any relationships that may exist between various assessment cells, any actions must be passed using the established Chain of Command.

The data collection plan may also initiate the process of appropriate military staff engaging with non-military organisations prior to the commencement of the operation.

A suggested guide to the type of information that should be published in Annex OO is provided in annex F.

Stage 3. Data Collection and Treatment

3.1 Data Collection

As part of operations assessment planning, and in addition to the data collection plan, consideration must be given to a number of factors that affect the way in which data is handled by the operations assessment staff once it has been passed to them from the designated data collectors.

The KD process provides a shared knowledge base of operationally-relevant material. Both the products from and the raw data used by the operations assessment process add to the understanding of the operational environment and different elements of this information will be both drawn from and feed back into the knowledge base. The KD and operations assessment processes are interdependent by the virtue of their common linkages to the knowledge base.

The operations assessment staff must consider the procedures, expertise and manpower required to deal with:

- **Format Conversion** – Data may be received in a number of standard formats, such as military signals or pre-defined questionnaires, or free formats such as newspaper reports or information posted on web pages. This data needs to be sorted and transposed into a format that is useable for analysis purposes, ideally into the knowledge base as a single location for data storage and management.
- **Data Storage** – Where possible, a single, centralised data storage mechanism makes data backup and therefore integrity easier, aids consistency by ensuring that different data users make use of the same sources, and through appropriate organisation and formatting as listed above allows automated searching. The precise method implemented in an HQ is heavily dependent upon the requirements of the KD staff, but a single database would provide the best functionality and flexibility, as opposed to, for example, a collection of spreadsheets, documents or presentations.
- **Data Relationships** – Just as the operational environment can be described as a system of systems, there is a hierarchical relationship between different data elements that

are used to determine the system states of that system of systems. These relationships can be maintained in both the way the data is organised and stored and in the metadata that accompanies it. Again the use of a database that provides relational database functionality, rather than a collection of individual files, enables this multi-dimensional organisation of data.

- **Security Classification** – The security classification is specific to each individual data type, and possibly each data element, depending upon both its source and sensitivity. The ability to note these individual classifications alongside the data is essential, especially if data is to be exchanged with external organisations.
- **Searchability** – The use of a single, centralised data storage capability enables complex but fast querying and searching of the data. The format for data storage, as mentioned above, is important to enable this (for example, storing dates in a standard format that enables searching by specific date, day of week, week, month etc.). During the format conversion stage of data collection it is important to add metadata or categorisation data to the raw data to enable this (for example, an incident report in Afghanistan may have lat/long positional data, but categorisation data for province and/or region could be added at the time of format conversion).
- **Presentation** – The way in which the user interacts with the data needs to be considered, depending upon the range and capabilities of the likely end users. These can vary from an experienced database user who can be trusted with direct access to the stored data, through to making data available through a web interface for wide ranging multi-user access and possibly a means of data entry. This issue needs to be considered as part of operations assessment planning.

3.2 Use of Data Collected

It is key to recognise that the data called for in the Data Collection Plan may not be immediately usable in an operations assessment process. A process of data-mapping – taking the raw data collected and putting it into the metrics – is generally required. For example,

the information sought for an MOE might be “*proportion of refugees without shelter*”, while the actual data collected might be:

- Total number of refugees.
- Total number of tents provided.
- Number of refugees per tent.

This requires calculation to determine first how many refugees have shelter (*total number of tents provided X number of refugees per tent*), then determine the proportion without shelter (*[total number of refugees – refugees who have shelter] / total number of refugees*). This data-mapping process may be very time-intensive early in the assessment process, but should take significantly less time once the ‘data map’, or process to move from raw data to metrics, is created.

It is also valuable to test the ‘data map’ prior to commencing the operation. This may also lend insight to the sensitivity of individual data streams or identify additional requirements.

It is also worth noting that individual raw data points may be used on multiple occasions in different combinations in order to create the required metrics (e.g. a raw list of incidents can be used to create metrics such as incidents per day, incidents by province, incidents initiated by own forces etc.). A centralised data store and management process ensures that data is only collected once, data is consistent wherever and whenever it is used in the assessment process, and updates to data easily and readily flows into updates to all of the affected metrics leading to potential updates of the overall assessment. The means to achieve this can be outlined in the operations assessment plan, and specified in detail in conjunction with the data collection plan.

3.3 Duplication of Reporting and Data Exchange

Assignment of units or resources to data collection must include a detailed check to ensure that:

- Data is gathered at the appropriate level (for example a tactical unit should not be responsible for collecting data that extends beyond their geographical or military area of responsibility and expertise).

- The same data should not be collected by multiple units. Two possible exceptions are:
 - 1) Data corroboration may be achieved by assigning multiple units to collect the similar data when those units are focused on different geographical regions, areas of influence, etc.
 - 2) Data validation using multiple sources may require multiple units to collect the same data.

Whenever possible, include required formats for data exchange to ensure interoperability. This speeds the analysis process and avoids confusion as data is passed through the Chain of Command. Data formats and reporting mechanisms should be promulgated from higher level Commands prior to the start of operations, if no format standards already exist.

Furthermore, it is evident that data collection using non-military organisations would be simplified if those organisations use common standards, formats and language. It may be up to the operations assessment staff to ensure that this particular request reaches the non-military parties involved by either working through their own contacts or through a CIMIC division or organisation, for example.

External organisations (either within the military command chain or non-NATO actors) that are being used to provide data may be encouraged by ensuring that they are given access to the wider data available, and in particular the analysis that it is used to create. This could assist both in ensuring data consistency, by allowing examination of historical data, and also motivation through demonstration that the action of collecting data that may not be immediately relevant to the data collector is being used to add value to the planning and execution of the overall mission.

3.4 Data Treatment

In an operational environment it is likely that data is gathered from a range of sources, and within each source individuals, processes and opinions will change, all of which affects the consistency and quality of the data supplied. Data can be considered of suitable quality if they are fit for their intended uses, whether that is short-term decision-making or longer term planning, or if they correctly

represent the real-world situation, which they are intended to describe.

Data should be checked using a 5-step process:

- a. Data Profiling – inspect the data for obvious errors, inconsistencies, redundancies or incomplete information.
- b. Data Quality – correct, standardise and verify the data, paying particular attention to and questioning those values that lie outside the expected range.
- c. Data Integration – match, merge or link data from a variety of disparate sources, looking deeper where independent sources provide different pictures.
- d. Data Augmentation – enhance data using information from internal or external sources that were not included in the original data collection plan.
- e. Data Monitoring – look at the longer-term history of data to ensure control of data integrity over time.

3.4.1 Procedures for Absence of Data

Absence of data may be caused either by random errors (e.g. a failure in the reporting chain on a particular day) or more systemic problems (e.g. the data is more difficult to collect than anticipated during the planning process). The potential influence of a missing data item depends on its position and relative importance within the analysis hierarchy. Where it is one of many metrics used to assess a particular system, and the instances of missing data are random, then it may be possible to simply overlook that data item for an assessment cycle. However, when the converse applies, that there are few metrics, or that missing data is systemic, then action needs to be taken to, in the short term, estimate the missing value, and in the longer term or systemic situation, change the metric being used. Where synthetic data has been used this must be clearly noted in both the assessments produced and within the data, so that any influence can be taken into account should it later be used to support other assessments or used for historical analysis.

Replacing the missing data item with a representative value rather than simply ignoring the issue is the most attractive option, as it minimises the disruption to the analysis process and, if correctly done, avoids the development of artificial bias in the assessment

results. However, correct replacement of the data is a complex problem beyond the scope of this handbook, and should be tackled by a suitably qualified and experienced analyst, who can use a range of statistical techniques to ensure that values are realistic and representative of the situation and the changes are achieved without distorting estimates or distributions where applicable.

3.5 Other Applications of the Assessment Data

The aim of operations assessment is to inform the Commander's decision making process with regards to progress against the OPLAN, and the foundation of this decision making is the data or evidence that is collected, analysed and assessed. Whilst operations assessment products will be created and managed by the operations assessment staff, there are likely to be many other functions throughout the HQ that are producing products for consumption both within and outside the HQ, and for which a common source of data as evidence will assist in the ensuring the HQ provides a consistent and coordinated view of the operation. Such uses may include:

- Facts to support Public Affairs
- Products to support Strategic Communications (STRATCOM) and Info Ops
- Evidence to support legal arguments
- Special assessments at the request of the Command Group
- Feeding and supporting analysis in the Lessons Learned process
- Presentations and papers for internal and external audiences
- Briefings to visitors
- 'Hot Facts' that can be quoted as evidence as and when required to support statements and arguments
- Indicating change in the operational environment which may make it necessary to engage different decision makers
- Making the case for assignment of forces by nations or other requests for resources (e.g. Crisis Urgent Requirements)
- Providing a historical perspective and reference point

- Continuity of understanding of an operation

In order to achieve this wide ranging exploitation of the data it will be necessary for a wide range of HQ staff to know of and understand its existence and utility. operations assessment staff will need to ensure that the data and its potential uses are visible and accessible across the HQ, through mechanisms such as a close working relationship with KD and Information and Knowledge Management (IKM), internal briefings, publishing on appropriate web pages, engagement with potential users, and production and dissemination of example products.

Stage 4. Analysis, Interpretation and Recommendations

4.1 Analysis

Once collected, data must be analysed so that valid conclusions can be drawn about the metrics. Change in these metrics must then be analysed in aggregate to determine progress towards individual effects or directly towards objectives. The analysis of metrics should form the main body of evidence brought forward to the final assessment. In preparing the assessment, analysed data is synthesised with other material such as expert opinion and commentary, in such a way that the Commander and other stakeholders can gain appropriate understanding of the current situation and make recommendations for future action.

Whilst the initial exploration and analysis of data is predominantly methodical in nature, the final development of the assessment is as much an “art” as it is a science.

The main responsibility for drawing together data and analysis lies with the operations assessment branch, and the process is normally set out in a HQ Standard Operating Procedure (SOP) or other instruction in order to synchronise the operations assessment effort with other aspects of planning and execution. However, it is important that there is dialogue between the operations assessment branch and other HQ staff branches, and with other actors providing input to the assessment process, in order to reconcile any inconsistencies and lead to the development of a consensus view on key issues wherever possible. In some cases aspects of the operations assessment process will be carried out by specialists outside the HQ, whether through “reach back” to higher HQs, NATO or national agencies, civilian subject matter experts or in collaboration with other actors.

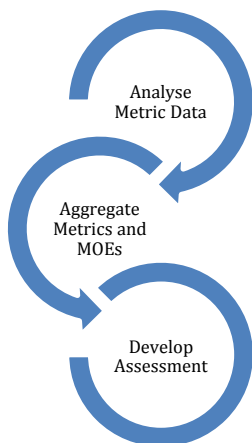


Figure 4. Analysis, Aggregation and Assessment

4.1.1 Data Analysis

Initial data analysis is the first examination of data collected in order to give early insight to any important features it possesses. The most important feature of new data in operations assessment is often whether it represents a change or not in the supported metric. In some cases this initial analysis only requires simple comparison of new data with that collected in the previous period. In the majority of cases, however, more advanced techniques to visualise, transform and process the data is required so that valid insights can be drawn.

Data analysis overlaps with and follows on from Stage 3: Data Collection and Treatment. Initial considerations should include:

- Categorising or grouping data in some way. Grouping and categorisation should already have been considered as part of the data collection plan. However, this may be reviewed at this stage for re-categorisation or re-grouping. This may be necessary if data collected has not been collected according to plan. For example, data collected on a daily basis may be more appropriately grouped into weekly totals, or other data collected at unit level summarised on a geographical or regional basis.

- Data may require corroboration or validation from additional sources. Care must be taken not to corroborate subjective evidence with objective evidence drawn from the same source.
- Data may require transformation. Transformation is different from grouping/categorisation in that it makes the original data into a different type. An example is taking a qualitative assessment reported by a subordinate unit and assigning it a category (e.g. High/Medium/Low) or a numerical score or range. Transformation may be required in order to allow for the data to be used by advanced analytical tools in subsequent analysis (e.g. for use in geospatial mapping and analysis software).
- Data may benefit from visualisation through charting or other graphical methods. A simple scatter plot of new data compared to previous data may reveal significant or unusual features. Where data has a geospatial aspect, simply plotting data on an appropriate map display may reveal interesting spatial aspects such as “hot spots.”

4.1.2 Identification of Change and Trends

A trend is a consistent trajectory of change over a number of intervals, or between several groups exhibiting the same kind of change. In order to identify a trend, it is important to identify whether a reported or apparent change is actually significant. “Significant” means that change has actually occurred and is not in fact due to some other factor, such as normal random variation or simple measurement error.

For some quantitative data it is possible to apply statistical tests to indicate whether a change over time or between groups (correlation) is significant. These statistical tests must be applied with care as they rely on certain assumptions about the structure of underlying data. Such statistical tests are usually expressed in terms of probabilities or “confidence” that a particular data point represents a change from a previous data point or a different source.

In other situations the method used to gather data is itself imprecise, and may also be associated with margins of error. For example, if an opinion poll survey is used to gather a particular piece of data, the way in which the poll is designed, and the size and makeup of the population sample polled, all give rise to uncertainty over the extent to

which observed change represents real change. Analysis of data arising from these methods should always acknowledge the error margin (usually expressed in a percentage) built in to the method (See annex E on Polling).

Where data is particularly “noisy” or subject to short term variability further consideration should be given to grouping, smoothing or curve fitting the data. This often makes it easier to help show trends in data and captures changes in the direction of the trend better than an unadjusted data.

Variability that can be predicted, due to seasonal weather conditions for example, should be built into the expectation of how a metric will progress during the planning process.

As an example, consider a metric ‘*The number of refugees passing a border crossing point*’, where the acceptable condition is indicated by 75 or fewer refugees per day. As shown in Figure 5 if the number is calculated on a daily basis there may be large fluctuations in the value, while smoothing the data over weekly or monthly time periods makes it easier to pick out longer term trends without provoking continual action and reaction based on the highly variable daily figures. Observing that the data is, in fact, highly variable might also lead to a review of the existing acceptable condition and its amendment to weekly or monthly-based values.

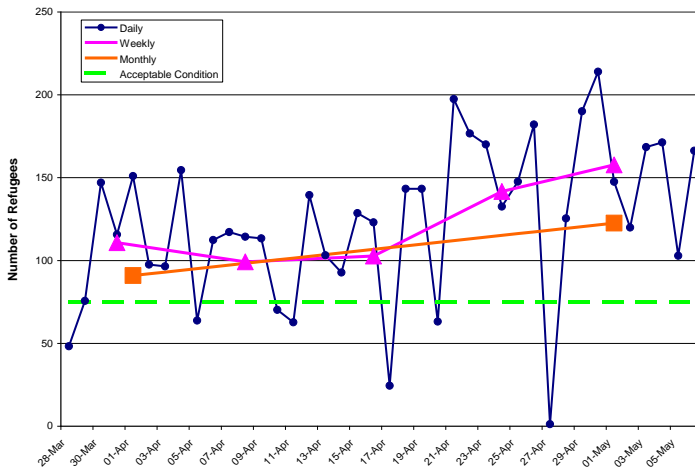


Figure 5. Data Smoothing Example

Another important technique that may be utilised is regression analysis. Regression analysis can be used to determine whether metrics are correlated – i.e. that they depend on each other in some way (interdependency). Understanding correlation between metrics is important particularly when it has been assumed in the data collection plan that they are independent. In some cases correlation – especially with delay – can be confused with causation. Causal links cannot be determined from correlation alone (see Section 2.1 for more detail on causality).

There are many other data analysis techniques available for quantitative and geospatial data and these should be implemented when appropriate, drawing on the skills and advice available within the operations assessment branch, elsewhere in the HQ and potentially through other specialist resources elsewhere in NATO. The application of advanced analytical techniques must not increase the complexity of an assessment. The sophistication of analytical tools applied towards an assessment should be transparent to the audience. Findings from advanced analytical techniques should always be presented clearly and in an easily understood manner.

Where the data collection plan specifies a DROC for a specific metric, the observed AROC must be analysed to determine whether or not it

matches the expectation of the plan. Additionally, it is also important to review any predictive aspects of change, using PROC.

Whilst there are a wide range of established techniques for quantitative data analysis, qualitative data analysis also requires a methodical approach. Change reported in narrative text from within the chain of command may overemphasise indicators of positive change and neglect similarly important indicators of lack of progress. Data originating from outside the chain of command may not be given due evidential weight, or its shortcomings may be opaque.

As noted above, it may be useful to transform qualitative data of this type into categories or linear scales. The concept of desired, actual and predicted rates of change may be more difficult to determine with qualitative data, even if categorised.

What is common in analysing data is that care must be taken in its analysis to ensure that the appropriate amount and quality of relevant data, corroborated where possible, is brought forward in order to justify each conclusion that significant change has occurred.

At the conclusion of the initial data analysis phase, operations assessment staff should have a clear idea which areas of data reliably indicate change, and which areas will be the focus of subsequent work in preparation for the assessment.

4.1.3 Aggregating Analysis

The operations assessment plan envisages that analysing the various metrics together should permit the operations assessment staff to conclude whether and how progress is being made towards the creation of desired effects and objectives.

Where an effect or objective is described by a single metric the direction of change (if any) in the effect/objective is determined by the change in the underlying metric. However, when there are multiple metrics underpinning an effect, it is important to decide what are the relative contributions made by each metric to the change (if any) in the overall effect. Mechanistic rules or weighting can rarely be justified, as it is likely that if multiple metrics are required then this will be because each metric is indicating the status of quite different sub-systems, and those differences mean that the state of the overall system cannot be derived by mathematically combining those metrics into a single meaningful mathematical value.

In all cases expert judgement is needed to determine what the appropriate aggregate change should be. The assumptions or rationale underpinning such judgement should form part of the final assessment. In some cases, change in an underlying metric may be considered to have made little or no contribution to higher level objectives. However, if an observed change in an underlying metric is consistently considered to have had no impact at higher levels, this should lead to review of the operations assessment plan.

4.1.4 Archiving

Data and its analysis may have uses far beyond the immediate needs of the assessment at hand. This may include short notice ad hoc questions arising from the assessment (e.g. “Is there a link between Factor A and Factor B?”) or for subsequent longer term studies. The data archive used for operations assessment should ultimately form part of the historical archive of the mission. In order to do this it is necessary not only to store that data as collected, but to be able to store enough information for analysts who are not familiar with the data to understand the reasons, context and methods that were the basis of its collection and use. This archiving function should be carried out in conjunction with the KD staff within the HQ (see section 3.1).

4.2 Reporting and Data/Result Visualisation

An important part of developing an assessment is reporting and presenting its key aspects. This section concentrates on good practice in portraying quantitative time series data, as this is likely to form a large part of most assessments. Although the precise choice of visualisation method used depends on the type of data to be displayed and the preferences of the Commander, the following points should always be considered:

- Does the visualisation selected portray the data and analysis as simply and clearly as possible?
- Is the conclusion to be drawn from the analysis logically supported by the visualisation?
- Does the visualisation have sufficient contextual information (labels, legends, sources) to be interpreted appropriately without too much clutter?

The examples illustrated in Figures 6 to 9 all focus on a notional typical effect: *Security in the region is restored to pre-crisis levels*. In this example this effect has a single associated quantitative MOE: *Number of attacks on populace by adversary forces*. Derived from the operational plan, it has been deduced that the principal own and allied force activities which contribute to this MOE are the ‘Number of own-force patrols’ and ‘Number of Host Nation patrols’, which can be combined to determine the *total number of patrols*. These values give a set of MOPs associated with the example MOE. With this example in mind a number of visualisation techniques are demonstrated in the following sections using the same underlying data.

The bar chart visualisation (Figure 6) simply shows how the recorded value of the MOE has changed from day 1 to day 19. Also indicated are the AC (in green) and both the Actual and the Predicted RoC (red and purple dashed lines).

In this case the PROC is based on a simple linear extrapolation of the previous 10 days values and the visualisation needs additional contextual information added to describe why this particular forecast had been selected over all others. Similarly, it is helpful to explain the rationale behind selecting the AC at that level; for example it could be based on a measurement or inferred from historical data of pre-intervention attack levels.

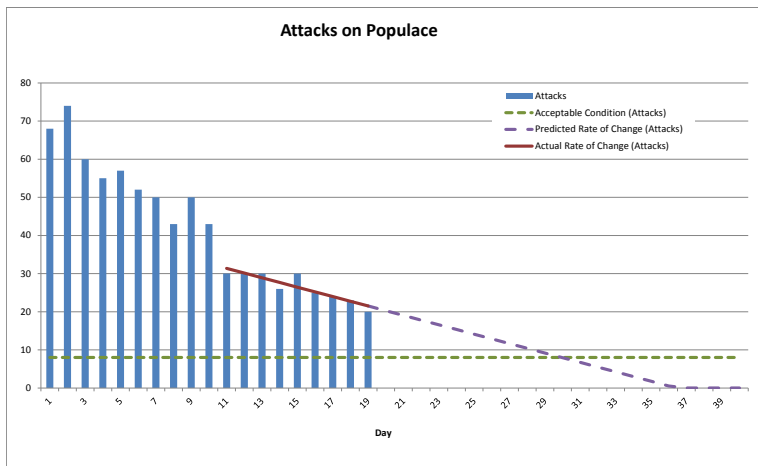


Figure 6. Simple bar chart showing No. of attacks on populace by adversary forces

Additional information can be overlaid onto the bar chart. Figure 7 adds three MOPs to the previous data; that for the number of daily Own Force patrols, the number of Host Nation (HN) patrols and the total number of patrols, including its PROC. Although the addition of the new data does appear to show a slight correlation between the increasing number of own/HN force patrols and the decline in adversary attacks, it is far from a definitive result. Moreover, the addition of new information arguably makes the visualisation more cluttered and confused.

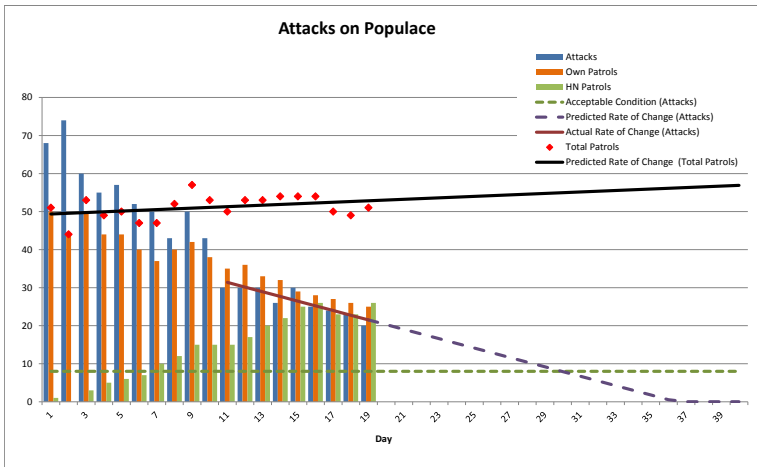


Figure 7. Cluttered bar chart with multiple metrics (MOE & MOP)

Figure 8 shows the same data as in Figure 7 but plotted as a line graph rather than individual bars. This makes some aspects of the data trends more apparent (such as the levelling off in number of HN patrols around day 15) when compared to the bar graph. However, it is still not apparent from the legend or other marginal information what insights are drawn from comparing these MOP and MOE time-series.

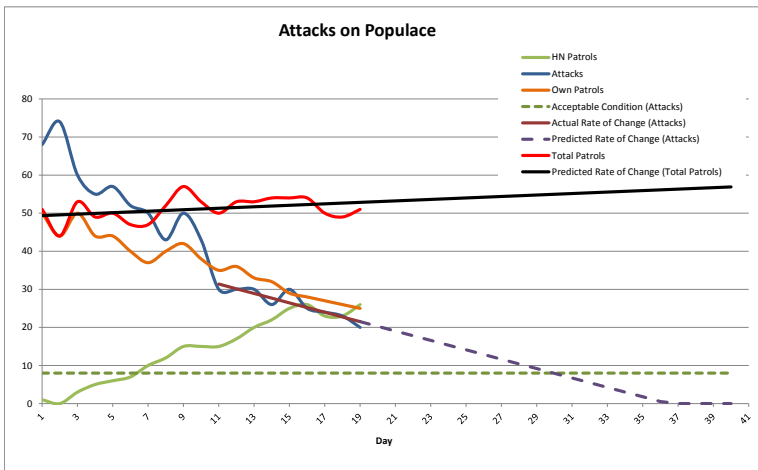


Figure 8. Line graph with multiple metrics - conclusion difficult to draw

Slider bars as demonstrated in Figure 9 reduce the mass of data available into a set of simple linear scales. Each slider bar represents the range of possible outcomes for the metric depicted, with desirable shaded in green and undesirable in red. In this example, the number of attacks potentially has no upper limit (i.e. it is an unbounded metric), so the choice of scale for plotting the points on the bar is arbitrary, and may cause problems if the metric value goes beyond the initially chosen limits at some point in the operation. For each slider the current observed value is plotted using the blue arrow on the lower side of the slider, which shows the change against the previous reporting period plotted in black on the upper side of each slider.

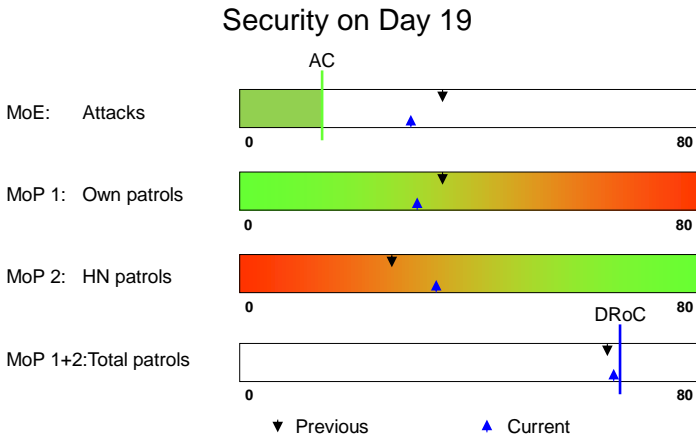


Figure 9. Slide Bars

Slider bars have some advantages in distilling a large number of metrics into a much smaller number of graphics, which can draw attention of senior leaders to key issues. However, they run the risk of oversimplifying trend information (it is difficult to assimilate change over more than 2 periods) and obscuring how any subordinate metric changes contribute to higher level progress. Caution should be taken when it comes to representing scale and associated colours on a slider bar to avoid implying that every metric resides on a similar scale (See section 5.2.1 on colour coding).

Geospatial Displays may be a powerful method of visualising data incorporating a spatial element. Spatial plots may reveal geographic insights such as hotspots or clusters of activity close to population centres or other features such as in Figure 10. Data need not have a precise location to be usefully visualised in a spatial way, when a shaded thematic map or choropleth may be used. Care should be taken not to overemphasise the values associated with larger areas (which naturally draw the attention of the viewer) at the expense of smaller areas; one way of correcting this is to normalise or scale the visualisation according to area as in Figure 11 which portrays the global distribution of AIDS cases in 2008 through shading (percentage of adults infected) and scaling of region size (absolute numbers).

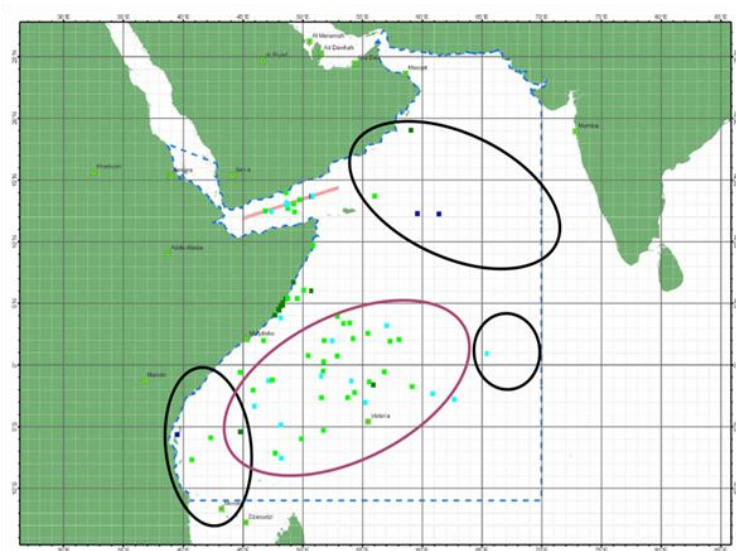


Figure 10. Geospatial Display

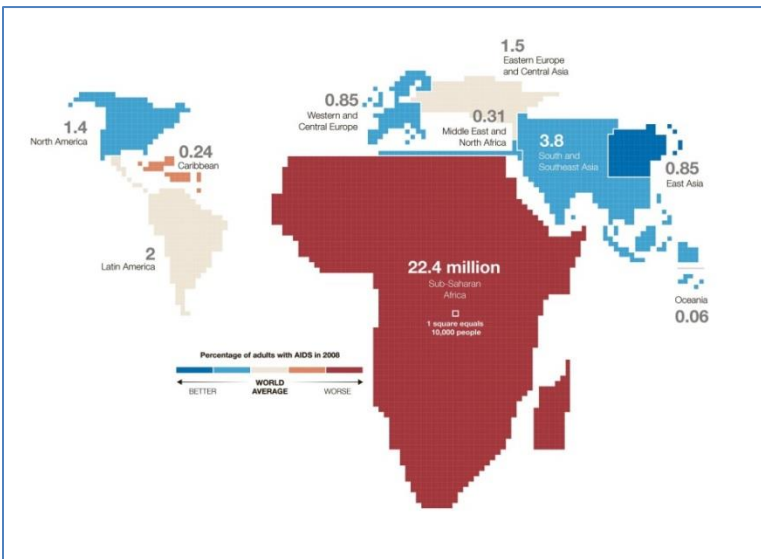


Figure 11. Choropleth: Global Distribution of AIDS 2008

Visualisation of qualitative data can also play an important role in clarifying key issues both for analysts and the intended target audience; basic principles of clarity and simplicity remain valid. Care should be exercised in simply using a graphic or visualisation for its own sake. Where non-numeric data has been transformed onto one or more linear scales, this may suggest an appropriate form for a chart or graphic.

Two examples of visualising qualitative and quantitative data in a radar chart are demonstrated in Figure 12. The first radar chart shows a quarterly assessment on which subjective ratings have been given on a linear scale of 1 (negative) to 10 (positive) for eight different MOEs. The second radar chart shows the results from a quarterly survey where public perception is used for eight different MOEs. The survey data is plotted on a scale of 0% to 100%, where each MOE is represented on one spoke of the radar chart. In addition, the radar chart clearly shows the acceptable condition of 65% plotted on the chart in green shaded area. One limitation of radar charts is that it is often difficult to plot changes over time on one chart without exposing the audience to a confusing plot. However, a radar chart can be replicated side by side to indicate changes over time, where the change in plots shape indicates the change between assessments.

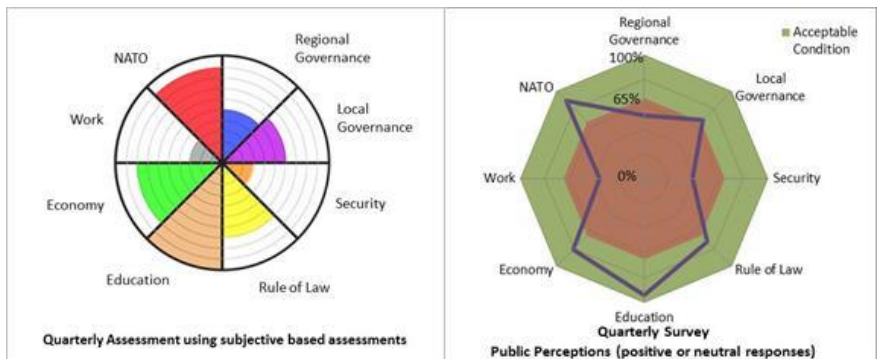


Figure 12. Examples of qualitative and quantitative data on Radar Charts

4.2.1 The Use of Colour Coding

Traffic light colours can provide a simple, easily comprehensible overview of the status of a particular factor, and are particularly

appropriate where an immediate and reactive decision must be made, e.g. an enemy attack is imminent - which weapon system(s) are available to counter it? There are significant drawbacks, however, to the use of traffic light colours in a deliberate decision making process where there is some time to interpret and digest a range of evidence in order to make confident decisions that may have a deeper and more profound impact on the course of an operation. These drawbacks include:

- Categorisation of colour boundaries is often highly subjective, with risk that the categorisation is based on the individual's feeling about a particular conclusion, rather than the most accurate objective representation. Defining categories of colour boundaries in assessment planning stage is important to reduce this risk.
- If categories have not been defined in advance a subjective categorisation of the analyst maybe further compounded by the subjective interpretation of the reader; i.e. an individual presented with a colour for a particular factor applies their own subjective interpretation of what that colour means, which is unlikely to match exactly what the originator of the colour intended⁹.
- There is likely to be a strong association between the assignment of colour and the acceptance of a level of risk¹⁰, i.e. where the individual assigning the colour feels risk to the mission is high they may choose red or perhaps amber. In the execution of an operation acceptance and ownership of risk is most definitely the domain of the Commander, and therefore the determination of risk associated with a particular decision should be the Commander's, based on the best evidence available.

⁹ As a simple example of the possibility of mismatch of interpretation between individuals, consider a metric that is expected to show a low level of achievement at the start of an operation, and a high level at the end. Person A may expect a traffic light system to show red at the start of an operation, amber in the middle of the operation as progress is made, and green towards the end of the operation. Person B may expect a traffic light to show green at the start of an operation when achievement is low but in line with expectations, green in the middle of the operation if progress is on track with what was planned, and green at the end of the operation when success is achieved.

¹⁰ Note: this is also dependent on different cultural interpretations .

- When aggregating quantitative data into a smaller number of qualitative categories such as, 'red', 'amber' and 'green', summarising into a single category loses the fidelity to understand changes over time (for example it may be critical to distinguish between an indicator that is amber but almost 'green' or another that is amber but almost 'red'). As an example, Fig 13 shows quantitative data of three different metrics plotted on three graphs. For each metric the situation remains 'amber' throughout time. If the qualitative assessment of simply 'amber' is given to the Commander for each assessment, then there may be a risk that the fidelity of changes in the data are missed. These graphs show an underlying time history with three very different trends which could possibly demand a different level of command attention. It is important that if a simplified assessment of 'Amber' is presented to the Commander, that the assessment staff have a deeper understanding of the trends over time in order to advise the Commander when necessary.

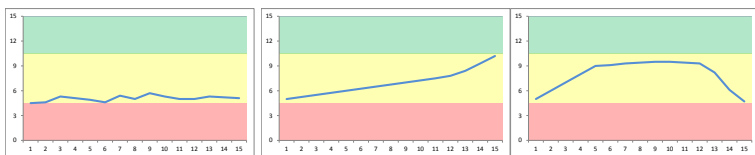


Fig 13. Colour coding versus underlying time history examples.

- There will be significant reliance on the initial assessment; if an issue is categorised as green it is likely to miss out on further attention, and so is highly dependent on the initial categorisation being reliable and robust.
- There may be pressure to report in line with perception. If a metric is colour coded in a way that mis-matches with the underlying perception of the situation, there may be discussion on the re-setting of colour boundaries, whereas the presentation of objective, quantitative data that is misaligned with perception is likely to lead to the more important and relevant discussion as to why the mismatch exists. If colour boundaries were to be adjusted to match perception then all objectivity in the assessment would be lost, and any evidence based assessment process that suffered this fate could be abandoned in favour of simple gut feel.

- A Commander's attention will probably be drawn more to those metrics that have been colour coded red rather than green. This is only valid—and avoids misdirecting attention from higher priority issues—if a colour change boundary can be objectively defined. This means that the importance of a colour change in any one metric has exactly the same impact, meaning, influence and importance to the mission, as a colour change boundary on any other metric—a situation that is unlikely. Traffic lights at best provide only a limited indication of the current state of a situation, and then only if all observers use the same interpretation of the meaning of the colours. They do not indicate if progress is being made.
- As substantial relevant information is hidden in the use of traffic light colours, they require many caveats and explanations to make them reliable and informative, especially in high risk critical decision making. A more detailed visualisation of the relevant data is likely to be a better and simpler solution overall.

4.3 Revalidation of Assumptions, Metrics, and Targets

4.3.1 Operations Assessment of Effect vs Action

Planning and operations assessment staff make assumptions during plan development, which are reflected in the operations assessment design. These assumptions are generally of the form “If *A* takes place then *B* will change.” One of the key roles of the operations assessment process is to check these assumptions wherever possible, in particular by assessing whether own force activity is making a contribution to achievement of effects. Four broad cases may be defined:

Case 1: Effect On Track, Action On Track

In this instance, the planned Effect is being created, is on track to be achieved when expected, and all planned actions are being accomplished. On face of it, this means that the assumptions underpinning effect achievement appear to be validated. However, it is worth considering that something other than the planned actions may actually be achieving the effect.

Case 2: Effect On Track, Action Not On Track

In this instance, the MOE indicates that the Effect is on track to be achieved, but the corresponding MOP indicates that some required Actions are not being accomplished as planned. Possible reasons may include:

- Invalid choice of MOP: the real status of the action is not being captured.
- Invalid assumptions regarding Actions to create Effect: something in the Engagement Space has created a change, but not the Actions.
- Invalid allocation of resources: less Action than planned was required to create the Effect. This may also be a poor choice of MOP goals.
- Invalid choice of MOE: the real system state is not being captured

Case 3: Effect Not On Track, Action On Track

In this instance, although planned actions are being accomplished, the desired Effect is not on track to be achieved. Possible reasons may include:

- Invalid assumptions regarding Actions required to create the Effect: not all of the correct actions are happening because the chosen Actions do not lead to the desired Effect.
- Presence of an unknown actor: the influence of an unknown or unexpected actor(s) is causing a negative influence on the system element, and is inhibiting the influence caused by own Actions
- Own force Actions are creating unintended and/or undesired effects that are counteracting the intended, desired effect
- Invalid choice of MOE: the correct system element is not being measured. This may also be a poor choice of MOE goals.
- The MOP goals may be insufficient and not accurately measuring accomplishment.
- An unforeseen time lag exists: conditions for the Effect have been created, but the Effect is not apparent yet.

Case 4: Effect Not On Track, Action Not On Track

In this instance, the desired Effect is not being achieved, and the planned actions are not being accomplished. Possible reasons include, amongst others, those listed in Cases 2 and 3.

In all cases the operations assessment staff must make a detailed analysis of the situation and recommend possible courses of action to the planning and execution staff. In some cases, revision of the operations assessment plan may be required.

4.3.2 Assessment of Other Planning Elements

The same assessment process can be applied to the higher level plan elements (e.g. Effects achieved, Objective not achieved). Again, without assuming causality it is possible in a qualitative judgement to compare the status of planning elements within the hierarchy of the operational design.

Comparison of the 4 cases as listed above will determine whether the creation of that system state is achieved or on track, or that changes to the plan are required.

4.3.3 Unintended Effects

An important part of operations assessment is investigation into unintended or unexpected consequences. Where unintended or unexpected effects threaten the achievement of objectives they may become risks to mission success.

4.3.4 Goals

Goals (see 1.9 Setting Goals) may shift as phases of the operation change, or as ongoing analysis of an existing MOE or MOP demonstrates that the initial goal is not accurate or when changing priorities or phases make a different goal more appropriate. In any case, the Commander must always approve changes to goals.

4.4 Art of Assessment: The complete assessment

The art of assessment describes the synthesis of analysed data with other information to form a credible assessment. Other information could include narrative or other assessments from subordinates, or expert opinion from inside or beyond the HQ. It is difficult to be

prescriptive about the art of assessment but the following principles should apply:

- No single piece of data should be given undue weighting or prominence unless there are clear and compelling reasons why this should be.
- The assessment should not draw excessive conclusions from limited or uncertain data, and should not claim that change has taken place if there is no evidence supporting that conclusion.
- Where analysis of available evidence contrasts with, or contradicts other assessments (for example, from a subordinate, a reputable NGO, media outlet or think tank) then it is important not simply to ignore the contradictory assessment but to consider why the contradiction exists and where possible, to reconcile these differences in the final assessment.

The format and content for any operations assessment report is determined by the situation and intended audience, and usually is specified in local SOPs. As stated earlier in the handbook, the purpose of operations assessment is to inform the Commander if the operation is being executed as planned and if the operation is achieving the desired results. The products from operations assessment provide advice directly to the Commander and recommendations for adjustments to the overall plan. The following key items should be included in the operations assessment:

- Statement of where the mission is placed overall in relation to its main objectives
- Key changes since the last assessment/divergences from the plan and supporting data and analysis
- Expert deductions from the analysis
- Predictions, where possible
- Conclusions and recommendations
- Proposed changes to assessment plan and any requirements for additional studies

Regardless of the format chosen, there are great gains in efficiency and collaboration if all levels of the command or organisation use similar or compatible report formats. Finally, there must be procedures in

place to determine report approval authority, required distribution, and authorized distribution.

4.5 Reasons to Reconsider the Operations Assessment Plan

The operations assessment plan is an integral part of the overall operational plan. As such, if anything in the operational plan changes, the operations assessment plan must be verified to ensure it remains current. Possible reasons for change may include, but are not limited to:

- Changes in political guidance or command direction.
- Unplanned or unforeseen changes in the Engagement Space.
- Transitions between phases of the operation.
- Changes in the validity of the planning assumptions.
- Reassessment of strategic or operational risks.

Methods of data collection and sources of data are part of the operations assessment plan; certain assumptions are made about the availability of data during the planning process. During execution of the operations assessment process, problems may arise due to the below factors, which are all reasons to revisit the operations assessment plan, and may lead to selection of different data collection techniques or different MOP/MOE:

- Poor quality data.
- Unreliable data or unreliable/inconsistent collection.
- Data appears to be or is proved inaccurate.
- Gaps appear in data, perhaps caused by non-availability.

A selection of tips and best practices for operations assessment, some of which may lead to review of the assessment plan, are listed in annex C.

Annex A. Operations Assessment in NATO

A.1. Operations Assessment at the Military Strategic Level

At the military strategic level, the term “Operations Assessment” refers to the development and conduct of the measurement of strategic progress and results of the post-North Atlantic Council (NAC) execution directive activities.

In the complex, multi-dimensional and asymmetric military operations of today and of the future, “success” is becoming increasingly hard to define. In previous years, the battle-damage assessment paradigm focused on military targets: numbers of enemy killed, bridges destroyed, or quantifiable metrics about the status of enemy military forces. Experience demonstrates that many extra factors must now be considered, as winning militarily may not necessarily lead to success in every domain.

At both the strategic and military strategic levels, the engagement space must be examined from a comprehensive perspective, across all PMESII domains, to ensure that all relevant influences, actors and interdependencies have been considered. Activity in the military domain affects – and is affected by – the activity and situation in the non-military domain. Operations assessment at the military strategic level must therefore assess progress in the non-military domains in addition to military progress and results. There may be many factors outside the military domain that are required to successfully achieve the NATO end-state, therefore a military success does not necessarily result in NATO success. A comprehensive military strategic-level operations assessment may identify recommendations that are beyond NATO’s mandate and can be raised at the NAC for national consideration.

At the military strategic level of command, operations assessment is therefore a function that involves varying combinations of: continuous measurement of system states to determine the achievement of strategic effects and progress towards the achievement of strategic objectives in both military and non-military domains; measurement of strategic progress, with consideration of results of activities of non-military organisations as they contribute to NATO objectives; an overall evaluation of progress towards the NATO end-state; and the

subsequent development of conclusions and recommendations that support Supreme Allied Commander Europe's (SACEUR) strategic decision-making, and inform the NAC.

A.1.1. Organisation, Roles and Responsibilities at the Military Strategic Level.

At the military strategic level, the Supreme Headquarters Allied Powers Europe (SHAPE) Comprehensive Crisis and Operations Management Centre (CCOMC) has the overall responsibility for the military strategic operations assessment. The CCOMC may seek outside expertise for certain aspects of the operations assessment function. At the military strategic level, operations assessment staff in the CCOMC have the following specific responsibilities¹¹:

- Considering the operational level operations assessments received from the operational level Headquarters and other areas of NATO, to produce the military strategic level operations assessments on ongoing military operations for SACEUR
- Producing for SACEUR the military strategic level operations assessments on all other domains, with inputs received from various internal and external sources
- Producing the operations assessments required¹² for the NATO HQ level.

As operations assessment at the military strategic level considers political, economic, social and infrastructure issues, the practice of operations assessment may be enhanced by the use of SHAPE's staff and available subject matter experts (SMEs) to better define and analyse the non-military aspects of a system. Operations assessment staff may seek experts from all relevant domains from: NATO organisations, including: the Intelligence Fusion Centre (IFC) or Civil-Military Planning and Support (CMPS) Section at NATO HQ; or non-NATO organisations, including: academia, think-tanks, international organisations, or private contractors.

¹¹ For each operation, duties and responsibilities may be shared and exchanged between levels, which will be defined in the operations assessment OPLAN annex.

¹² The normal strategic operations assessment product required by the NATO Crisis Management Process (NCMP) is the Periodic Mission Review (PMR).

Operations assessment at the military strategic level often requires data, frequently openly available, from international organisations such as the United Nations, World Bank, Organisation for Economic Cooperation and Development, European Union, Organisation for Security and Cooperation in Europe, International Monetary Fund and the Red Cross. All these organisations have well-developed Monitoring and Evaluation¹³ capabilities, and have detailed reports and subject matter expertise on many conflict areas (See annex B for more detail on civilian organisations). At the military strategic level authorisation and guidance on practicalities of how assessment staff should exchange assessments with these organisations should be directed in the OPLAN or at a later stage, as relationships develop, by Fragmentary Order (FRAGO).

A.1.2. Characteristics of Operations Assessment at the Military Strategic Level

In NATO, planning is initiated as a top-down process that begins with the NAC issuing a decision sheet tasking the NATO Military Authorities to provide an assessment of the crisis situation. In response to the NAC decision sheet and the associated tasking, SHAPE will produce a SACEUR's Strategic Assessment, informing the NAC decision process and eventually serving as a baseline assessment for operations planning (Phase 2 of the strategic Operations Planning Process (OPP) as described in Chapter 3 of the COPD).

Once the decision has been made to initiate strategic planning, planners will begin developing the strategic Operational Plan (OPLAN) (Phase 4a/4b of the strategic Operational Planning Process). Operations assessment staff will develop the design of the operations assessment which includes metrics to measure progress and effectiveness.

The strategic OPLAN considers strategic military effects and objectives that contribute to achieving the NATO end-state, in combination with non-military effects and objectives. The design of the strategic operations assessment describes the means by which progress will be measured towards the creation of military strategic effects and achievement of military strategic objectives, and their

¹³ Monitoring and Evaluation is the equivalent term to "Operations Assessment" that is generally used by international organisations.

contribution towards setting the conditions necessary to achieve the NATO end state, as well as progress in the various non-military PMESII domains in the engagement space. The development of this design should commence during the initial phases of planning. It contributes to the process of defining system state changes and actions by ensuring that these can indeed be observed and measured. Furthermore, the process of determining metrics increases understanding of the corresponding effects and objectives.

Success cannot be defined in military terms alone. A comprehensive operations assessment of the strategic engagement space and the progress towards the NATO end-state must consider all the aspects of the PMESII domains within the region and the engagement space. Although NATO does not have the instruments of power to act directly in many of the PMESII domains, a comprehensive operations assessment at the military strategic level must consider all domains within the region and the engagement space. These aspects include:

- Progress and effectiveness of NATO military operations.
- Development of political processes, governance, and civil institutions.
- Security, rule of law and human rights.
- Economic development.
- General well-being of local populations.

Military strategic level operations assessments may be produced for a variety of different purposes and audiences:

Primary Audience	Focus of Operations Assessment	Intended Use
SHAPE, NAC, Military Committee, Host Nations	Comprehensive operations assessment in all PMESII domains	Briefing the NAC Informing NATO political decision-making and strategic communications

Primary Audience	Focus of Operations Assessment	Intended Use
SHAPE	Strategic overview of ongoing military operations	High level decision-making and necessary adjustments (within the scope of the current strategic OPLAN) Inform military strategic level operations assessments and decision-making processes
Operational HQs	Strategic implications of progress and effectiveness of operational level missions	Decision-making on necessary plan adjustments for the operational level

At the military strategic level, the primary type of metric will typically be MOEs, however there will be a requirement for a small number of military strategic MOPs (e.g. access to strategic basing or conducting regional engagements).

The military strategic operations assessment design will incorporate the use of three sets of metrics:

- A set that measures the creation or achievement of the elements in the military strategic plan, which contribute to the achievement of the NATO end state, military strategic effects and objectives.
- When required, a set received from the operational level, that measures creation of effects, establishment of DCs, and achievement of objectives and performance of actions in the operational OPLAN, some of which are directly linked to strategic elements.
- An independent set that may not be directly tied to elements in the strategic plan, but considers the broader PMESII aspects of the engagement space. The progress towards military strategic objectives, for example, will not always be revealed by an amalgamation of MOEs from the operational level. These independent MOE may capture standard data produced by international organisations.

In general, operations assessment at the military strategic level will consider longer timescales than operational and tactical levels. Currently, NATO produces a Periodic Mission Review (PMR), which is the formal operations assessment of strategic progress and results normally required by the NATO Crisis Management Process (NCMP). Depending on the specific context and situation, the timescales may change, or different strategic operations assessment products will be required.

A.1.3. Summary – Operations Assessment at the Military Strategic Level.

Operations assessment at the military strategic level is much more than a simple aggregation of lower level operations assessments. The strategic engagement space is a complex, interdependent system of systems including: regional and international powers and political institutions, regional, national and international economies, social and cultural influences, international organisations and non-governmental organisations, humanitarian aid organisations, reconstruction and development agencies, and military forces, both NATO and national.

NATO's instruments of power are military and political; however, SACEUR also requires an understanding of how NATO military operations interact with non-military domains, how the activities of non-military organisations contribute to or hinder progress towards achieving military strategic objectives and contributing to achievement of the NATO end state, and how the values of various critical social and economic indicators change.

Operations assessment at the military strategic level not only focuses on the overall progress of NATO military operations and the general state of critical PMESII domains, but also considers relevant non-NATO actors. If cooperative planning is conducted with specific non-NATO organisations, cooperative operations assessment should occur.

In some cases, it may be necessary for strategic level operations assessment to take an expanded view and consider two separate missions where interdependencies exist between the two operations. As an example, the humanitarian assistance mission in Pakistan and the ongoing International Security Assistance Force (ISAF) mission in 2005, where the former operation, if properly synchronized and coordinated with the NATO mission in Afghanistan, could have had positive strategic impact on the latter.

A.2. Operations Assessment at the Operational and Tactical Level

The primary focus at the operational and tactical levels of command is the execution of the military operation and the creation of effects, the establishment of DCs and the achievement of the operational objectives defined in the plan. The operation is planned by the Operational Planning Group (OPG) and assessed by the Assessment Working Group (AWG).

Plans will need continual adjustment, based on the circumstances of the operation, to be effective. The primary purpose of operations assessment at the operational and tactical levels is to increase the effectiveness of the execution of military operations. By continually monitoring and analysing the implementation of actions, creation of effects, establishment of DCs and achievement of objectives, the intention of operations assessment is to assist the Commander in making evidence-based adjustments to the plan being executed. Operations assessment aims to provide confirmation of the plan design, by demonstrating that the planned actions are indeed creating the desired results, and to improve understanding of the workings of the engagement space. Operations assessment also plays an important role in providing situational awareness relative to the plan.

At the operational level, the process is based on the overall analysis of metrics measuring progress of planned actions or DCs (MOP), the creation of desired effects, the establishment of planned DCs and the achievement of planned objectives (MOE). These two aspects, although closely related, are the subject of two separate assessment activities.

Medium-Term Assessment seeks to answer the question “*Is the operation being executed as planned?*”. It specifically measures progress in the generation of desired effects towards the achievement of Decisive Conditions (DC) as stipulated in the operational level Operational Plan (OPLAN). The medium-term operations assessment is an integral part of the Commander’s Decision Cycle, and is performed regularly (i.e. once per Battle Rhythm). It consists of a mid-term review of current actions and their desired effects leading to DC along particular Lines of Operation (LOO). It assumes that the approved operational design is correct and therefore measures the progress in generating the planned effects toward the achievement of the planned DC. It uses a combination of MOP and MOE to generate

the necessary data and evidence to formulate the assessment. The main output of this assessment are recommendations to the Commander for the adjustment of own actions aimed at generating the effects more efficiently and effectively and indirectly the achievement of the DC more efficiently and effectively. It could also result in recommendations to adjust the effects leading to DCs. The changes resulting from this type of assessment will generally be promulgated through a Fragmentary Order (FRAGO) or a Joint Coordination Order (JCO).

Long-Term Assessment seeks to answer the question “*Is the operation achieving the desired result?*”. It specifically measures progress towards the achievement of the objectives. The long-term assessment is conducted at longer intervals (over several medium-term assessment cycles) to allow enough time for the operational environment to evolve sufficiently following the creation of planned effects and to draw meaningful conclusions about the adequacy of the plan to achieve the objectives. The long-term operations assessment consists of a long-term review of the achievement of the operational objectives, and seeks to answer the questions:

- If the DC are being achieved but the operational objectives are not, how should the operation be adjusted?
- What influences, if any, outside of the military domain, are limiting or preventing achievement of the operational objectives?
- Are the military strategic objectives and the operational objectives achievable within an acceptable timeframe?

This involves the monitoring of the broader environment across all PMESII domains where they impact on the military operation. It uses metrics specifically related to the operational objectives to generate the necessary data and evidence to formulate the assessment. The main outputs of this assessment are recommendations to the Commander for changes such as the adjustment of the existing operational design (LOO, DC, effects etc.) and requests to the strategic level to influence other factors or actors outside of the military domain which are preventing the effective and efficient achievement of the operational objectives. The change resulting from this type of assessment are generally promulgated through a JCO or OPLAN revision.

At the tactical level, the focus is on measuring the achievement of planned actions, tasks or activities using MOP. In some special cases, the tactical level may measure the establishment of DCs and creation of operational effects using MOE.

A.2.1. Organisation, Roles and Responsibilities at the Operational and Tactical Level

At the operational level, the Commander owns the operational level operations assessment. The operations assessment staff takes responsibility for development of the operations assessment annex in the OPLAN (Annex OO), and the conduct of operations assessments during execution. At the operational level, operations assessment staff have the following specific responsibilities:

- Acting as the focal point for operations assessment development in their respective HQ, including the preparation of operational level operations assessment related lessons as a contribution to doctrine development.
- Working with the OPG during development and revision of the OPLAN.
- Considering the tactical level operations assessments received from their subordinate commands and other areas of NATO.
- Producing the operational level operations assessments on on-going military operations considering the tactical level operations assessments.
- Contributing to military strategic operations assessments, as required.
- Monitoring the operational level risks.

At the Tactical level, the Commander owns the tactical level operations assessment. The operations assessment staff takes responsibility for development of the operations assessment annex in the OPLAN, if required, and the conduct of operations assessments during execution. At the tactical level, operations assessment staff have the following specific responsibilities:

- Acting as the focal point for operations assessment development in their respective HQ, including the preparation of tactical level operations assessment related lessons as a contribution to doctrine development.

- Working with the OPG during development and revision of the OPLAN.
- Considering the operations assessments received from their subordinate commands and other areas of NATO.
- Producing the tactical level operations assessments on on-going military operations considering the operations assessments of their subordinate commands.
- Contributing to operational level operations assessments as required.

A.2.2. Operations Assessment Process at the Operational and Tactical Level

It is essential that operations assessment personnel are involved from the beginning of the decision cycle¹⁴ of plan, execute, monitor, and assess to ensure that the plan is measureable.

Operations assessment staff are an integral part of the OPG and support the planning in the different syndicates. The syndicate developing the operational framework must contain operations assessment expertise. The operational framework consists of operational objectives nested within the Military Strategic Objectives (MSO), related operational effects and DCs. The operational design¹⁵ is the key reference for the plan and operations assessment process, and thus forms the basis for the development of the operations assessment annex.

In order to achieve an overall coherent operations assessment plan, the operations assessment development must be conducted as a top down approach throughout all levels of command. Consequently, the operations assessment products at military strategic level, especially the military strategic objectives and effects, and the military strategic operations assessment design must be taken into consideration at the operational level.

Both the planning process and the development of operations assessment products are interdependent. They both must be derived

¹⁴ See AJP-01(D) paragraph 0524

¹⁵ The Operational Design is the fusion of the operational framework and the Commander's initial intent

from the operational design. It should be a key goal of the operations assessment staff to develop the operations assessment annex in parallel whilst the OPG finalises the rest of the OPLAN.

When the main body of the operational OPLAN is drafted, the operations assessment annex must be developed using the expertise of all members of the OPG (See annex F for more – Annex OO). The development of MOEs and MOPs must be coordinated with the relevant SMEs and subordinate commands to ensure coherence and validity of the assessment plan.

Assessment Working Group. During execution, periodic meetings of the AWG ensure that the plan is on the correct track or identify potential plan adjustments for submission to the Commander. The frequency of an AWG will be dependent on local SOPs and Operational Tempo.

The roles and responsibilities of the AWG include:

- Ensuring data collection is occurring as planned
- Monitor achievement of objectives, DCs, effects and actions
- Develop assessments
- Achieve consensus across the HQ
- Draft assessment products for approval

The AWG must have an interdisciplinary make-up in order to maintain coherence within the HQ. AWG membership should include:

- Assessment Staff
- Data contributors
- Planning staff
- Special Advisors
- External Liaison or assessment staff from subordinate commands

Outputs of the AWG may include:

- Periodic assessment products as required by Commander and/or Higher or other HQs
- Support to the assessment processes of higher HQs

- Analysis of the assessment inputs from subordinates
- Monitoring of risks, unintended or unforeseen effects
- Development of Annex OO to the OPORD/OPLAN
- Monitoring and review of the suitability of MOEs, MOPs
- Monitoring the data collection process

Beyond the AWG, interactions with J2/KD provide key data and analysis for the operations assessment staff. In turn, the operations assessment staff provides feedback to systems analysis and KD to help ensure a common perspective.

Assessment Board. The AWG will provide the appropriate data for the Assessment Board (AB) briefing to the Commander. The AB is the formal forum to seek Commander's endorsement of the operations assessment provided. The AB should culminate in a recommendation to the Commander.

The roles and responsibilities of the AB include:

- Agreement upon a common understanding of the state of an operation
- Assessment products synchronised with the requirements of higher HQs
- Allow for Commander's direction and guidance on moving the operation forward
- Approval of assessment products for dissemination outside the HQ or outside NATO.

Similar to the AWG, the AB must have an interdisciplinary make-up. However the membership is different and should include:

- Command Group
- Special Advisors
- Division Heads
- Liaison Officers

Outputs of the AB should include:

- Approved assessment products for distribution such as routine operations assessment report, Periodic Mission Review, Operational Updates.
- Direction and Guidance issued to other boards and working groups (See NATO's Allied Command Operations Directive 80-70 for more detail on campaign synchronisation)

The resulting Commander's decision and direction will initiate staff actions and plan adjustments (e.g. FRAGO, Joint Coordination Order (AD 80-70 Campaign Synchronisation and Targeting), development of branches and sequels, plan revision) and adjustments of the operations assessment Annex OO if required.

A.2.3. Summary – Operations Assessment at the Operational and Tactical Levels

It is essential to recognise that operations assessments at all levels are not isolated, but need to be considered in a holistic way in order to understand the whole theatre of operations and beyond. Operations assessment is done to monitor and validate the plan during execution and can be a significant part of the decision-making process. Without operations assessment, decision makers will find it more difficult to get the appropriate feedback (plan-execute-monitor-assess).

The operational level is the pivotal point in the overall coherent NATO operations assessment process, as it acts as the interface between the strategic/political requirements and tactical operations.

A common understanding of operations assessment requirements and procedures throughout all levels of command is to be achieved and continuously maintained via appropriate operations assessment liaison structure, information exchange, meetings and exercises. Operations assessment is a HQ responsibility.

A.3. Interrelations between Levels of Command

A.3.1. NATO Headquarters and SHAPE

SHAPE is responsible for providing military advice to NATO Headquarters with regard to the planning and execution of operations. Past and likely future progress can be drawn from the operations

assessment process, and integrated together with outputs from planning and other processes. The conclusions and recommendations are fed to NATO Headquarters in formal products such as Periodic Mission Reviews and Operational Updates.

Whilst operations assessments at the operational level will look at the achievement of military objectives, the strategic military assessment produced by SHAPE may provide military advice on progress towards the NATO end state. The military strategic level assessment is also likely to look in more detail at the non-military PMESII domains and their influence on the military operation, and also consider the activities of non-military actors.

Within NATO Headquarters the political decision making process will aim to gain consensus on the military advice by the Military Committee supported by the International Military Staff, and then consensus in the North Atlantic Council based on further work undertaken in committees such as the Operations Policy Committee supported by the International Staff.

A.3.2. SHAPE and Operational Headquarters.

The military strategic level initiates the overall operations assessment process as a top-down approach and gives guidance to the operational level regarding structure of the plan and reporting procedures. The operational level, as the pivotal point in the overall coherent NATO operations assessment process, requires that guidance from the military strategic level in order to ensure consistency. Clear reporting guidance from the Strategic Command supports the operational Commander's reporting requirements. In order to maximise collaborative work, military strategic and operational levels must ensure that their planning and operations assessment staff are fully integrated.

A.3.3. Operational and Tactical Headquarters

During planning, liaison or planning, experts of the tactical commands support the OPG and ensure the synchronisation of planning efforts between the levels of command. The operational design and the operations assessment annex will be the leading references for tactical level planning and operations assessment.

A.4. Summary of Operations Assessment at the Military Strategic, Operational and Tactical Level

The following tables provide a quick summary overview of the different operations assessment considerations, both military and non-military that the assessment at each level should address. In addition, the tables provide the differences between the audience or intended user of the operations assessment for each level.

Level	Operations Assessment - Considerations and Audiences			
	Military Considerations	Non-Military Considerations	Audience / Users	Geography
Military Strategic	<ul style="list-style-type: none"> Creation of the effects and objectives in strategic OPLAN, which contribute to achievement of NATO end state Progress of overall mission and status strategic military assets Capture of overall operations assessments from operational levels Engagement of internationally recognised subject-matter experts on region 	<ul style="list-style-type: none"> Achievement in political, economic, civil, social domains in theatre, as they relate to the achievement of NATO's aims. Achievements of key non-military national government, international, and non-governmental organisations, in theatre, as they relate to the achievement of NATO's aims. Tracking of international organisation's monitoring and evaluation in region (e.g. United Nations reports, World Bank, International Monetary Fund (IMF), Organisation for Security and Cooperation in Europe) Monitoring of key international conditions and situations that may impact upon strategic military mission (e.g. international trade embargos, world oil prices, international public opinion) 	<ul style="list-style-type: none"> SACEUR / SHAPE NAC NATO Nations' Defence Ministries Operational Level Commander Host Nation Government International Organisation (IO)/NGO HQs International Media 	<ul style="list-style-type: none"> International Regional Joint Operation Area (JOA)

Level	Operations Assessment - Considerations and Audiences			
	Military Considerations	Non-Military Considerations	Audience / Users	Geography
Operational	<ul style="list-style-type: none"> Creation of the effects, establishment of decisive conditions, and achievement of objectives in operational OPLAN Capture of operations assessments from subordinate level Coordination of overall data collection effort Hiring of external contractors required to support data collection / polling etc. 	<ul style="list-style-type: none"> Measurement of key conditions and situations in non-military domains that impact on the operational military mission Achievements of non-military organisations whose goals are specified in the military plan (either through collaborative planning or through estimation) 	<ul style="list-style-type: none"> Operational Level Commander Tactical Commander CCOMC Local IO / NGO partners Local host nation government Local and regional media 	<ul style="list-style-type: none"> Regional JOA

Level	Operations Assessment - Considerations and Audiences			
	Military Considerations	Non-Military Considerations	Audience / Users	Geography
Tactical	<ul style="list-style-type: none"> Establishment of decisive conditions as appropriate Achievement of tasks / tactical objectives / mission Data collection for the tactical / operational level operations assessments 	<ul style="list-style-type: none"> Data collection activities as assigned by higher commands Measurement of key conditions and situations in non-military domains that impact on the tactical level military mission 	<ul style="list-style-type: none"> Tactical Commander Operations Planning Group Local media Local Host Nation Commanders 	<ul style="list-style-type: none"> JOA

Annex B. Civilian Considerations in Operations Assessment

B.1. Assessment of non-military domains

Unlike in national contexts where operations assessments could be conducted in an inter-agency manner with civilian agencies providing assessment of non-military domains, operations assessments staff must conduct these assessments mainly using internal resources.

There are three main reasons why NATO operations assessment processes need to consider assessment of non-military domains:

1. NATO-led forces are responsible for providing support to non-military objective as part of NATO's contribution to the comprehensive approach to operations.
2. NATO-led forces are responsible for non-military tasks such as training and advising of local security forces in domains like rule of law, counter-corruption, transparency and accountability and gender equality.
3. As non-military aspects of the operations environment are considered key for the long-lasting success of military operations, assessment of these aspects will continue to be important.

While sharing assessments with civilian/ non-NATO organizations has been identified as important for NATO's assessment of non-military domains, operations assessment staff needs to develop understanding and processes for the assessments on non-military domains so, when sharing assessments with others, NATO operations assessment staff would know what to ask for and how to interpret the assessment received.

Most of the assessments that NATO could receive from civilian organizations will be in the non-military domains such as governance, development and regional cooperation, or political, economic, social and information issues.

Although the principles and methods of assessment might be the same for both military and non-military domains, the non-military domains have their own specificities.

- Non-military aspects pervade all domains of the security environment and assessment of security/ military aspects cannot be done without considering non-military aspects. Moreover, security/ military domain itself cannot be properly understood without consideration of the social, political, economic and information aspects.
- Despite the fact that there is an overall understanding that military and non-military aspects of the operations environment influence each-other and that long-term security and stability cannot be established without improvements in the non-military aspects of the operations environment, there is limited understanding on how these aspects influence each other and the overall success of a military operation.
- Non-military domains of the operations environment are multi-dimensional.
- Civilian organizations have limited involvement in the security sector and can provide limited assessment of non-military aspects of the security sector. Thus, NATO needs to develop its own mechanisms to assess non-military aspects of the security sector.
- Different organisations use different definitions of the same concepts and different operationalisations of these concepts, and NATO assessment staff needs to be aware of these differences when using external assessments or data to inform its own operations assessment.
- Different organisations use different methodologies that may not be available for the NATO operations assessment staff and this can result in errors in own assessment. For example, not being aware of the margins of error of an assessment done by another organization, operations assessment staff could interpret variations in numbers as changes while they could very well be within the margins of error.
- Among all different civilian organizations, there seems to be about 400 indicators of governance and more than 300 of development. These indicators could have been designed for certain purposes and could mislead NATO's own assessment if used in ways that depart from their original purpose or ignore statistical limitations of data.

- Civilian organizations could cite different indicators without adequate explanation of the problem they measure and the method of measurement.
- Civilian assessment may focus on specific indicators and neglect the broader picture of the operations environment and, using these indicators could create the danger of reductionism.
- Indicators measured by civilian organizations may be useful for describing the state of non-military domains but could fail to explain the causes and processes that led to that state.
- Some organisations tend to perform simplistic performance assessment focusing on output (e.g. delivery of food or services) but fail to capture behavioural and attitudinal changes that occur on long term.

B.2. Purpose and Scope of Considering Civilian Assessments

In any operation, there will most likely be non-hostile, non-military actors in the field such as government aid, development and diplomatic agencies, international organisations (IO), non-governmental organisations (NGO), and private commercial organisations. Many of these organisations conduct activities very similar to operations assessment; however, they generally use different terminologies and methods. For example, what NATO terms as ‘Operations Assessment’, is akin to the civilian practice of ‘Monitoring and Evaluation’ (M&E).

As part of the overall philosophy outlined for NATO’s contribution to a Comprehensive Approach, it is prudent for NATO operations assessment staff to gain awareness of Civilian Assessment terminologies and methods, in order that an exchange of operations assessment with external organisations can occur, and so that NATO Assessment staff can make use of civilian Assessment reports.

By exchanging and sharing assessments with relevant IOs, NATO will bring objectivity and comprehensiveness to assessment and it will allow NATO to address first hand with IOs any conflicting assessments. Additionally the benefits gained from IO’s assessments, information tools and the extended period they may have been

operating in-country would without a doubt improve the overall quality of NATO's assessments.

There are several key principles that underpin cooperation between organisations, and represent preconditions for successful exchanges of data and assessment with external actors:

- **A common language** – understanding of terms, understanding the purpose of assessments and evaluations, clarity on definitions and methodologies to prevent misunderstandings.
- **Development of trust** – research demonstrates that inter-organisational cooperation is facilitated by the existence of previous relationships and mutual trust and respect between cooperating parties. The earlier a relationship is developed with an organisation's evaluation staff, the better the chances for successful growth of trust. If left to initiation of a Crisis Response, the less likely the trust will be cultivated successfully.
- **Documentation handling and security classifications** – Documentation security and handling must be carefully agreed to ensure sensitive assessments from NATO and civilian organisations are not released into the public domain and trust destroyed. 'Confidentiality Agreements' may be set up for each organisation where required.
- **Mechanisms of interaction have to be simple and adaptable** – A spectrum of possible interaction levels with civilian organisations in operations assessment can include: shared training; comparison of conclusions on effects of common concern (e.g. security sector reform); co-ordination and coherence in common areas of operations assessment (e.g. surveys); establishment of in-theatre, joint civil-military operations assessment teams (where organisations have the same strategic aim or United Nations (UN) mandate as NATO).

B.3. Important Factors to Consider

While there are obvious benefits to exchanging assessment and evaluation, there are challenges associated with increasing cooperation. These challenges derive from the variation between

actors that exists despite commonalities. The exchange of assessments will occur on a spectrum of low to high cooperation; this level of exchange will vary with each situation and actor:

- **Differing goals:** Organisations have different goals and objectives; therefore, determining a common measure of success relative to those goals is challenging.
- **Political nature of Assessment:** Assessments are inherently political as they question the capability and operating assumptions of organisation, and affect how funding is allocated.
- **Transaction costs of cooperation:** Exchange involves using additional organisational resources, time and money, and depending on the circumstances, may reduce organisational independence. The following organisational costs may be present:
 - 1) Rules and communication structures for exchange of assessments must be created on a case-by-case basis with each organisation;
 - 2) Most organisational processes and structures differ in important respects, therefore resource effort must be expended to determine at what level to engage, with whom, and with what process to interact.
- **Differing contexts:** Although methodologies for assessment are similar amongst the communities, results, conclusions and recommendations will be contextual and tailored to each organisations' particular needs, therefore, conducting a synthesis using results only may be challenging.
- **Differing organisations:** Organisational structures, policies, practices, mandates, cultures, and missions vary significantly. Furthermore, the activities, structure and mandates of organisations will vary significantly with mission type (e.g. humanitarian relief, development, or post-conflict stabilisation).

B.4. Non-NATO Civilian Actors and the Civilian Sector

Many operations involve non-military actors working alongside military actors with varying degrees of coordination and cooperation;

operations assessment staff should make a concerted effort to learn more about the operating principles, missions and cultures of these non-military actors. This section describes some of the typical civilian sectors commonly encountered. Note, this section is not intended to be exhaustive, and is intended to give only an indication.

There is much commonality between the ways that the military community and the **international development community** approach analysis, planning, implementation and assessment of progress. International development work focuses on long-term and equitable economic growth, global health, agriculture and trade, democracy and governance, conflict management, humanitarian assistance, and many other factors. Creating and building sustainable, host-nation owned capabilities is a primary goal of much international development work. Organisations such as the World Bank, the United Nations Development Group, the European Commission and national development agencies, such as the US Agency for International Development, are informative places to look in order to learn more about the international development community. Further, many universities offer degree programs in international development.

Peace operations comprise peacekeeping – the provision of temporary post conflict security by internationally mandated forces – and peacebuilding – those efforts undertaken by the international community to help a war-torn society create a self-sustaining peace¹⁶. This multi-faceted community of interest includes military organisations, but tends to be oriented towards much more than just military principles and objectives. Some guiding principles of peacekeeping operations include consent of the parties to the conflict, impartiality in dealings with the parties to the conflict (not be confused with neutrality or inactivity), and non-use of force except in self-defence and defence of the mandate¹⁷. Other success factors include legitimacy, credibility and local ownership¹⁸.

Peace building has become an overarching term for an entire range of actions designed to contribute to building a culture of peace and can

¹⁶ The Henry L. Stimson Center in Washington DC, <http://www.stimson.org/fopo/programhome.cfm>

¹⁷ United Nations Peacekeeping Operations, Principles and Guidelines, 2008, adapted

¹⁸ Ibid.

include activities such as the promotion of a culture of justice, truth and reconciliation, capacity building and promotion of good governance, supporting reform of security and justice institutions and socioeconomic development¹⁹. For more information, see the documents footnoted below and their originating organisation's websites (www.un.org/Depts/dpko/dpko and www.oecd.org/dac). Also, organisations such as the US Institute of Peace (www.usip.org) and the Pearson Peacekeeping Centre (www.peaceoperations.org) publish extensively on the topic of Peace Operations.

Humanitarian assistance is the aid and action designed to save lives, alleviate suffering and maintain and protect human dignity during and in the aftermath of emergencies²⁰. The provision of humanitarian aid is driven by several fundamental principles including:

- Respecting and promoting the implementation of international humanitarian law, refugee law and human rights.
- Allocating humanitarian funding in proportion to needs.
- Ensuring, to the greatest possible extent, adequate involvement of beneficiaries in the design, implementation, monitoring and evaluation of humanitarian response.
- Strengthening the capacity of affected countries and local communities to prevent, prepare for, mitigate and respond to humanitarian crises, and
- Providing humanitarian assistance in ways that are supportive of recovery and long-term development²¹
- “Do No Harm”²²

More information can be found at the websites of the International Red Cross and Red Crescent (www.icrc.org), the UN Office for the Coordination of Humanitarian Affairs (www.ochaonline.un.org), international non-governmental organisations such as World Vision,

¹⁹ OECD/DAC Guidance for Evaluating Conflict Prevention and Peacebuilding Activities, 2007

²⁰ www.globalhumanitarianassistance.org

²¹ Principles and Good Practice of Humanitarian Donorship, endorsed at the International Meeting on Good Humanitarian Donorship, 17 June 2003.

²² Anderson, Mary B., Do No Harm: How Aid Can Support Peace - or War, 1999

and Oxfam International, and national government agencies such as the US Agency for International Development.

B.5. Civilian Sector and Operations Assessment

Within the humanitarian aid and development community, the function of Assessment is generally known by the term **monitoring and evaluation (M&E)**. Many of the major IOs, NGOs, and government agencies have well established M&E processes, and many have entire departments within their organisations to deal with this task. Although there is a broad spectrum of terminologies and processes within the civilian M&E community, there have been some attempts to standardise approaches by the Development Assistance Committee (DAC) of the Organisation for Economic Cooperation and Development (OECD).

The DAC is an international forum of 24 countries where donor governments and multilateral organisations – such as the World Bank and the United Nations – come together to help partner countries reduce poverty and achieve the Millennium Development Goals. The DAC issues analysis and guidance in key areas of development and forges ties with other policy communities to co-ordinate efforts. Its members also work together through peer review to assess each other's aid policies and their implementation so as to promote good practice. The DAC's objective is to be the definitive source of statistics on Official Development Assistance (ODA).

In 1991 the DAC set out broad principles for the evaluation process for DAC members²³. These principles were refined into five criteria that have been widely used in the evaluation of development initiatives – efficiency, effectiveness, impact, sustainability and relevance. Subsequently the criteria were adapted for evaluation of complex emergencies²⁴, becoming a set of seven criteria: relevance/appropriateness, connectedness, coherence, coverage, efficiency, effectiveness, and impact. The DAC criteria are intended to be a comprehensive and complementary set of measures.

²³ OECD-DAC (2000) DAC Criteria for Evaluating Development Assistance. Paris: OECD.

²⁴ OECD-DAC (1999) Guidance for Evaluating Humanitarian Assistance in Complex Emergencies. Paris: OECD.

Using the DAC frameworks, many major IOs and NGOs developed M&E frameworks. It is recommended that operations assessment staff working with civilian organisations become familiar with the OECD-DAC documents and the terminology employed. The OECD-DAC published a terminology guide, available on the website (www.oecd.org).

B.6. Civilian Approaches to Operations Assessment²⁵

Civilian research, literature and practice in the field with regard to M&E offer many lessons for military assessment practitioners. The following definitions are generally accepted by the civilian community for M&E²⁶:

Monitoring

A continuous function that uses a systematic collection of data on specified indicators to provide management and primary stakeholders of an intervention with information regarding the use of allocated funds, the extent of progress, the likely achievement of objectives and the obstacles that stand in the way of improved performance.

In general, monitoring is an activity that is used to answer questions such as:

- What happened?
- What is happening?
- When did it happen?
- Where did it happen?

Evaluation, on the other hand, tends to be a function that is more oriented towards answering questions such as:

- Why did things happen? / Why did things not happen?
- How did things happen?

²⁵ Adapted from USJFCOM (2010). Assessment of Progress in Military Operations: Considerations, Methodologies, and Capabilities, version 2.0. Produced as part of the MNE 6 concept development and experimentation campaign.

²⁶ Source: OECD www.oecd.org

Evaluation

The process of determining merit, worth or value of an activity, policy or program. It consists of the systematic and objective assessment of an on-going or completed project, program or policy, its design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, development efficiency, effectiveness, impact and sustainability. An evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision-making process of both recipients and donors.

Some similarities and differences between M&E are highlighted in table B-1²⁷:

Monitoring	Evaluation
Continuous or periodic	Episodic, ad hoc
Program objectives taken as given	Program objectives assessed in relation to higher-level goals or to the development problem to be solved
Pre-defined indicators of progress assumed to be appropriate	Validity and relevance of pre-defined indicators open to question
Tracks progress against a small number of pre-defined indicators	Deals with a wide range of issues
Focuses on intended results	Identifies both unintended and intended results
Quantitative and qualitative methods	Qualitative and quantitative methods
Data routinely collected	Multiple sources of data
Does not answer causal questions	Provides answers to causal questions
Usually an internal management function	Often done by external evaluators and often initiated by external agents

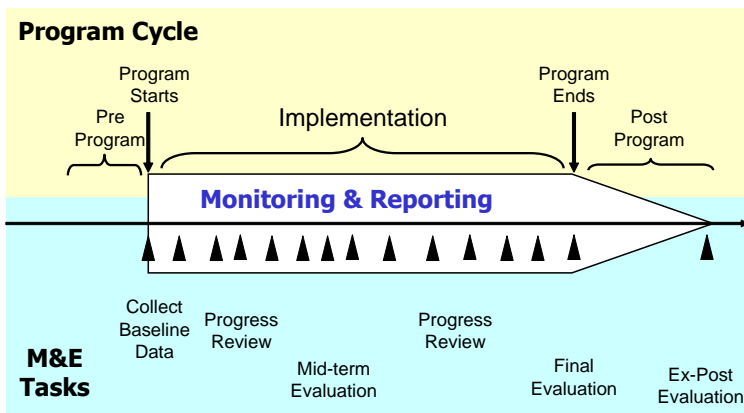
Table B-1 Monitoring versus Evaluation

²⁷ Source: Swedish International Development Agency, (2007) www.sida.se.

Key terms associated with M&E include:

- **Input:** The financial, human, and material resources used for the development intervention.
- **Activity:** Actions taken or work performed through which inputs, such as funds, technical assistance and other types of resources are mobilised to produce specific outputs.
- **Output:** The products, capital goods and services which result from a development intervention; may also include changes resulting from the intervention which are relevant to the achievement of outcomes.
- **Outcome:** The likely or achieved short-term and medium-term effects of an intervention's outputs.
- **Impact:** Positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended.²⁸

Figure B-1 Monitoring and Evaluation Tasks



Monitoring is normally an ongoing function, while evaluation is more episodic. The basic fundamentals of when M&E tasks should occur are shown in Figure B-1.

²⁸ Ibid.



Figure B-2 World Bank 10 Steps to a Results-Based Monitoring and Evaluation System

As demonstrated in the graphic above, monitoring and evaluation are both functions that are meant to be fully integrated into a design, planning, management/execution and assessment cycle. One example of this kind of end-to-end process model that may be of use for military application is the World Bank's 10 Steps to a Results-Based Monitoring and Evaluation System (2004). The overview graphic from that model is shown in Figure B-2 above:

There are many different types of monitoring and many different types of evaluation, all designed to suit specific needs. In addition to the documents that have already been referenced in previous recommendations, some documents, organisations and websites related to this topic include:

- Church, Cheyanne and Rogers, Mark, Designing for Results: Integrating Monitoring and Evaluation in Conflict Transformation Programs, 2006
- USAID, Monitoring and Evaluation in Post-Conflict Settings, 2006
- Social Impact Toolkit on Monitoring and Evaluating Fragile States and Peace building Programs
- World Bank Independent Evaluation Group, www.worldbank.org/oed
- OECD/DAC Network on Development Evaluation www.oecd.org
- USAID EvalWeb, www.evalweb.usaid.gov
- USAID Monitoring and Evaluation TIPS Documents

- Monitoring and Evaluation News, www.mande.co.uk
- The Measure Project, www.cpc.unc.edu/measure
- Collaborative Learning Projects, Reflecting on Peace Practice Project, www.conflictsensitivity.org/files/publib/Reflecting_on_Peace_Practice.pdf

B.7. Comparison of Civilian and Military Approaches

B.7.1. Origin of Results-Based Management and The use of “Effects” in Operations Planning²⁹

Perhaps somewhat unintentionally, the use of effects in operations planning and Assessment began under the same philosophy as results based management (RBM). The air-power military strategists that first developed “effects-based” thinking during and after the 1991 Gulf War had realised that by focusing on the consequences of destruction of Iraqi air-defence rather than the actual destruction, coalition planners were able dramatically to increase the air strike assets available³⁰. This led to the realisation that the way in which the accomplishment of objectives were assessed needed to be rethought, as focusing on percentage damage was misleading when infrastructure did not need to be damaged to produce a certain effect. What the military strategists had emulated was an RBM construct of planning, managing and assessing operations.

The RBM philosophy was first described in Drucker’s seminal management text *The Practice of Management*³¹, in which he

²⁹ Adapted from Williams, A. & Morris, J. (2009). Theory-driven evaluation in the military: Theory on the front line. *American Journal of Evaluation*, 30(1), 62-79.

³⁰ Where before, 10 bunker-busting bombs requiring several aircraft would be planned for to completely destroy an air defense station, the air operations staff soon realized that the effect of one single bomb was to cause the Iraqi operators to immediately shut down that station to prevent detection—thus, the same effect was created: Deny use of enemy air defense. However, the military “footprint” in strike craft and supporting logistics had been significantly reduced. See Deptula, D. (2001a). Effects-based operations: A change in the nature of warfare. Arlington, VA: Defence and Airpower Series, Aerospace Education Foundation; and, Deptula, D. (2001b). Firing for effects. *Air Force*, 84, 45-53.

³¹ Drucker, P. F. (1954). *The practice of management*. New York, NY: Harper and Row

delineates the process of management by objectives. His techniques were used by private sector management, before being adopted by the United States Agency for International Development which developed the concept of a “Logframe” or Logical Framework—an analytical tool used to plan, monitor, and evaluate projects³². The name arises from the logical linkages set out by the planners to connect a project’s means with its ends³³. The technique was adopted by many democratic governments in the 1980s as the philosophy of “New Public Management” emerged, which was driven in part, by a recognised need for governments to be more accountable and responsive to the public, and generally more efficient. In the civilian context, the core aim of using RBM is to shift the focus of planning, managing, and decision making, from inputs and processes, to the results to be met.

In a similar vein, the change from “traditional” military planning to effects in operations planning shifts the focus from inputs, methods and targets, to outcomes (effects), and furthermore, places an emphasis on effects at all levels of the campaign. The traditional military planning system considered results at the military strategic level, whereas at the operational level planning was focused on methods.

The developers of effects methodology realised that a military plan constructed from a list of treatments and outcomes, as specified in a log frame, did not provide the “operational art” gained by understanding the detailed causal relationships in a plan. Consequently, the development within NATO of the use of effects in operations planning was strongly integrated with the development of *systems thinking* methods and techniques. Similar issues with the RBM methodology are also extant in the civilian community. As such, much emphasis is placed on *theories of change* as an integral part of any evaluation approach.

³² Rogers, P. J. (2007). Theory-based evaluation: Reflection ten years on. In Mathison, S. (Ed.), *Enduring issues in evaluation: The 20th anniversary of the collaboration between NDE and AEA*. New Directions for Evaluation, Vol. 114. San Francisco: Jossey Bass; Sartorius, R. H. (1991). The logical framework approach to project design and management. *American Journal of Evaluation*, 12, 139-147.

³³ UNESCO. (2007). Results-based programming, management and monitoring (RBM) at UNESCO: Guiding principles. Bureau of Strategic Planning, UNESCO, Paris, France

A theory of change is simply an explanation of how and why a set of activities will bring about the changes a project's designers seek to achieve³⁴. A theory of change model describes the linkage between project inputs, the implementation strategy and the intended outputs and outcomes³⁵. A theory of change is not an academic hypothesis, but rather an everyday expectation about "how the world works"³⁶. Such theories can take the simple format: "We believe that by doing X (action) successfully, we will produce Y (movement towards peace)", or be much more complex³⁷. Theories of change are not statements of assumed linear causality, but rather a descriptive statement of planner's thoughts about how to change the environment. A useful first step in improving military strategies and plans would be for us to be more explicit about the "theories of change" that underpin our plans. The referenced documents include a much deeper description of theories of change and include detailed examples. The USAID Office of Conflict Management and Mitigation is currently (July 2009) conducting research related to developing and applying theories of change that may be of use to military planners as well. Finally, www.theoryofchange.org is a website "devoted to the application, methodology, use and practice of Theory of Change methodology"³⁸.

Table B-2 shows how NATO operations planning, execution and assessment can be considered as a type of results-based management system.

³⁴ Lederach, Neufeldt and Culbertson, *Reflective Peacebuilding: A Planning, Monitoring and Evaluation Toolkit*, The Joan B. Kroc Institute for International Peace Studies, University of Notre Dame, Copyright © 2007

³⁵ OECD/DAC Guidance for Evaluating Conflict Prevention and Peacebuilding Activities, 2007

³⁶ Lederach, Neufeldt and Culbertson, 2007

³⁷ OECD/DAC, 2007

³⁸ www.theoryofchange.org

Results-Based Management	NATO Operations Planning
Formulate objectives (results)	Define End State, Objectives, Decisive Conditions, Effects and Actions
Select indicators to measure progress towards each objective	Define Measures of Effectiveness (MOE) for each Effect, Decisive Condition, Objective, and End State. Define Measures of Performance (MOP) for each Action.
Set explicit targets for each indicator to judge performance	Establish goals, Acceptable Condition (AC), Rates of Change (ROC) for each MOE and MOP
Regularly collect data on results to monitor performance	Regularly collect data for MOE and MOP
Review, analyse, and report actual results vis-à-vis the targets	Conduct assessment of Effect, Decisive Condition and Objective status; conduct assessment of Action status to determine progress toward the target (End State)
Integrate evaluations to provide complementary performance information	Conduct assessment of Effect or Decisive Condition versus Action status; Effect versus Decisive Condition, and Decisive Condition or Effect versus Objective status to find insight as to why results were or were not achieved.
Use performance information for accountability, learning, and decision making.	Use the results of this assessment to inform the Commander and staff; identify required changes to both the plan and the assessment.

Table B-2, Comparison of RBM with NATO Operations Planning

B.7.2. Comparison of Civilian and NATO Terminology

Table B-3 provides a comparison of OECD DAC terminology with NATO terminology. Even within the civilian community there is a wide spectrum of terminologies and typologies used in monitoring and evaluation (assessment), however, the work performed by the OECD-

DAC has provided some standardisation. Table B-4 illustrates some comparisons between OECD-DAC and other civilian definitions.³⁹

³⁹ Adapted from Haugevik, K & Carvalho, B (2007). *Civil-Military Cooperation in Multinational and Interagency Operations*. Norsk Utenrikspolitisk Institutt (NUPI) 2007; and, Marthinussen, E. et. al. (2010). *Progress Assessment in a Multinational Operation – a Norwegian Perspective*. Paper presented at the 4th IMA Conference on Analysing Conflict Transformation, University of Oxford, June 2010.

OECD DAC definitions		Associated NATO (working) definitions	
Monitoring	A continuing function that uses systematic collection of data on specified indicators to provide management and the main stakeholders of an intervention with information regarding the use of allocated funds, the extent of progress, the likely achievement of objectives and the obstacles that stand in the way of improved performance.	Operations Assessment: The activity that enables the measurement of progress and results of operations in a military context, and the subsequent development of conclusions and recommendations that support decision-making. (Proposed definition to be ratified)	Operations Assessment
Evaluation	Evaluation refers to the process of determining merit, worth or value of an activity, policy or program. It consists in the systematic and objective assessment of an on-going or completed project, programme or policy, its design, implementation and results.	It should be noted that the term evaluation in a military context is often associated with an individual or unit performance evaluation as a step towards certification and should not be confused with the civilian view of evaluation.	
Indicator	Quantitative or qualitative factor or variable that provides a simple and reliable means to measure achievement, to reflect the changes connected to an intervention, or to help assess the performance of a development actor.	A metric used to measure a current system state.	Measure of Effectiveness (MOE)

OECD DAC definitions		Associated NATO (working) definitions	
Indicator	Quantitative or qualitative factor or variable that provides a simple and reliable means to measure achievement, to reflect the changes connected to an intervention, or to help assess the performance of a development actor.	A metric used to determine the accomplishment of actions.	Measure of Performance (MOP)
Goal	The higher order objective to which a development intervention is intended to contribute	The NAC statement of conditions that defines an acceptable concluding situation for NATO's involvement.	End state
Objectives	The intended physical, financial, institutional, social, environmental, or other results to which a project or programme is expected to contribute.	A clearly defined and attainable goal to be achieved.	Objectives
		A combination of circumstances, effects, or a specific key event, critical factor, or function that when achieved allows a Commander to gain a marked advantage over an opponent or contribute materially to achieving an operational objective	Decisive Condition

OECD DAC definitions		Associated NATO (working) definitions	
Impact	Positive or negative, primary and secondary effects produced by an intervention, directly or indirectly, intended or unintended. Results that lie beyond immediate outcomes or sphere of an intervention and influence the intensity, shape or likelihood of a conflict.	A change in the behavioural or physical state of a system (or system elements), that results from one or more actions, or other causes.	Effect
Outcome	The likely or achieved short-term and medium-term effects of an intervention's outputs.		
Output	The products, capital goods and services which result from a conflict prevention and peace building intervention.		
Activity	Actions taken or work performed through which inputs, such as funds, technical assistance and other types of resources are mobilised to produce specific outputs.	The process of engaging any instrument at each level in the in the engagement space in order to create (a) specific effect(s) in support of an objective.	Action
Input	The financial, human, and material resources used for the development intervention.	The available resources	Resources

Table B-3 OECD DAC and NATO Definitions Comparison

UN DP	EU	IASC	IFRC	WB	NATO
Intervention	Operation Response Mission	Response	Operation Response Mission	Operation Response	Campaign Operation
Activities Input Implementation	Activities Action Conduct	Activities Action Efforts Implementation	Activities	Activities Action Input Implementation	Action Decisive Condition
Impact	Initial Results	Benchmarks	Impact Benchmark	Impact	Effect Decisive Condition
Objective	Objective Target Goal Aim Purpose	Objective Mission Goal	Objectives Input Priorities	Objective	Objective

UN DP	EU	IASC	IFRC	WB	NATO
Result Output Outcome	-----	Result Output	Output Outcome	Result Output Outcome	End State
Monitoring Evaluation Assessment	Monitoring Evaluation Assessment	Monitoring (Real-time) evaluation Assessment	Monitoring Evaluation Assessment	----	Operations Assessment

Table B-4 Comparison of International Organisations Definitions

Annex C. Operations Assessment – Challenges, Key Messages and Top Tips⁴⁰

C.1. The Challenges of Operations Assessment

Military operations are characterised by a complex constellation of actors, organisations, priorities and pressures, which makes it difficult to conduct operations assessment that will enable senior leaders to answer the question: how are we progressing? In order for operations assessment to be relevant and useful in this context, it should meet three equally important but differing aims: 1) it must be adaptive to the operational tempo and ever-changing context of military operations; 2) it must make sense to the organisational configuration, process, and leadership and 3) it must be empirically supported and, ideally, objective. Yet there are a number of challenges that underlie and exacerbate the tensions between these three aims.

C.1.1. Institutional factors

There are several institutional challenges, the first of which lies in the bureaucratic structure of military organisations. Within NATO and other military settings, operations assessment is usually conducted in a command structure with multiple reporting lines and key stakeholders distributed across many organisations. Operations assessment is primarily conceived as a tool to support a Commander's decision making. Yet assessment products also drive decision making at subordinate command levels, and feed the assessment processes of higher level commands. An operations assessment can have multiple stakeholders at various levels, some of which have agendas that may not align with the needs of the Commander who owns a particular assessment. Finding a balance between measuring mission progress and meeting the various demands of other stakeholders, particularly at higher HQs, is a key challenge.

The second institutional challenge is the tension between the political imperative to demonstrate progress and the generally decades-long timescales over which change occurs in some complex operations such as counterinsurgencies. National political and budgetary agendas generally run on four or five year timescales and the political

⁴⁰ Adapted from draft UK JDN 2/12 Assessment and SAS-089 study.

narratives in the past have advanced the notion of quick and decisive military action, thus an alignment of political and military strategy to relatively short-term timescales is inevitable. In results-driven organisations there will always be subtle pressures to demonstrate success. Consequently, higher level assessments of progress may be framed so as to favour a more positive view than the reality on the ground.

The third institutional challenge concerns the way “progress” is politically defined and high level objectives are left intentionally vague. Operations assessments are set in the context of a political decision making process. In NATO each member must consider their own domestic political situations in addition to the political situation within NATO as a group. In military interventions there is rarely an objective reality to be empirically discovered, with high level goals such as “change,” “progress,” “freedom” or “security” treated as relative concepts politically constructed by various stakeholders. As a result, evidence is often contested depending on how issues are framed⁴¹. This can make it difficult for operations assessment and analytical personnel who strive to approach the task from a strictly objective and scientific perspective.

C.1.2. Requirement for comprehensive assessment

A comprehensive assessment of progress should include the entire range of factors in the operational environment, including many factors outside the security-domain. While military forces can report on the level of violence or security due to their privileged access to data and sensors, it is hard to make causal assessments on why the level of violence changed without a deeper knowledge of changes in other domains. For example, there is a known dependency between the level of violence and the extent of international aid and development in an area⁴², thus it is critical for military-focused operations assessment to consider all domains in order to understand key relationships.

⁴¹ Stone, D. (2002). *Policy Paradox: The Art of Political Decision Making*. New York: Norton Press; Fisher, F. & Forester, J. (1993). *The Argumentative Turn in Policy Analysis and Planning*. Durham, NC: Duke University Press.

⁴² Myerson, R. (2011). Rethinking the fundamentals of state-building. *Prism* 2(2):91-100. http://www.ndu.edu/press/lib/images/prism2-2/Prism_91-100_Myerson.pdf. Zyck, S.A. (2012). How to lose allies and finance your enemies: The economisation of conflict termination in Afghanistan. *Conflict, Security and Development*, 12(3):249-271.

The relationship between military and non-military actions and objectives changes between strategic and tactical levels, with non-military objectives defining the goals or objectives for military action at higher levels, while at tactical level it is more symbiotic. It is important for assessment to reflect this.

C.1.3. “Over-metricising”

There is a tendency for assessment staffs to overstate the number of metrics needed, thus generating a huge data collection requirement on subordinate organisations. A major challenge is for the assessors to understand “how much is enough” in terms of metrics and to appreciate the significant amounts of work needed to collect the data. The assessors should also give the data collectors an understanding and appreciation the value of their contributions⁴³.

C.1.4. Assessment in multi-actor environments

In an intervention there are often multiple local, national and international organisations undertaking a range of concurrent activities with different underlying objectives and goals over different time horizons⁴⁴. Each organisation has their own definitions of success and measures of progress. Yet often the success of one actor’s operation depends on the successes of other actors’ operations. To be truly comprehensive, assessments should be made both with inputs from, and consideration of, all salient stakeholders in the operational theatre. NATO employs civilian staff with relevant civil-domain expertise in crisis operations, and they can contribute to the operations assessment process. Non-military stakeholders may “frame” the intervention objectives in different terms than the military, which can result in multiple perspectives that are difficult to reconcile.

C.1.5. Data collection

There are many practical difficulties in data collection. Data collectors may face security threats resulting in limited or no access to key data

⁴³ Downes-Martin, S. (2011). Operations assessment in Afghanistan is broken. *Naval War College Review*, 64(4):103-125.

⁴⁴ Stolk, C et al. (2011). *Monitoring and Evaluation in Stabilisation Interventions: Reviewing the State of the Art and Suggesting Ways Forward*. Santa Monica, CA: RAND. http://www.rand.org/pubs/technical_reports/TR962.html

sources. Data collection by military-affiliated personnel may unduly influence responses by locals. Ensuring consistency in how data is collected over time can be difficult. Assessing attitudinal change in populations, especially in conflict areas, can be particularly challenging. Opinion polling or other survey methods can help monitor such attitudes, but polls and surveys are difficult to conduct properly. Institutional pressures and the challenges of war may also incentivise data collectors or operators to reduce quality controls, or in the worst cases manipulate data.

C.1.6. Establishing reliable baselines

Many military operations are complex and turbulent and it is often difficult to establish reliable baselines or to distinguish between strategic shifts in the environment and shorter-term fluctuations. Sometimes assessment designs and data collection systems are not developed in advance of an operation, resulting in evolving data collection and assessment approaches. Even if there is the capacity within a command chain to handle a large number of metrics, setting appropriate targets for those metrics is far from simplistic. Defining an acceptable condition to be reached by a metric requires an understanding of what is locally and contextually “normal.” An understanding of what is “normality” may change with time or as the underlying premises and assumptions of missions change. Furthermore, the focus of missions may change over time, meaning the importance of metrics vary over time. Some commands may stop collecting data for some metrics, only to realise months or years later that the lost data would have been useful for current priorities.

C.1.7. Methodological challenges

There are several key methodological challenges that face operations assessment policy makers and practitioners. First, there is the difficulty in deciding when metrics are best measured using quantitative or qualitative techniques. Second, there is the challenge of data quality and reliability—whether quantitative or qualitative.

Third, there is the challenge of determining the appropriate level of aggregation. Goals are often described as an aggregate of lower level effects to be achieved, each of which is associated with one or more metrics. The aggregation of the subordinate metrics into a higher level metric is far from trivial. The aggregation of data into a composite

measure may conceal some important features of the situation. Nevertheless the challenge of aggregation remains, as it is a natural human tendency to seek simplified explanations for complex situations, especially when time pressures preclude consideration of a broad range of inputs. Filtering down to only the key information is a critical assessment skill.

C.1.8. Use of assessments by decision makers

Finally, there is the basic problem of how evidence is used by decision makers. Assessments are just one out of a multitude of inputs received by a Commander. Social science literature offers a range of explanations on how policy makers use data and analysis to support their decisions⁴⁵ and the varied uses of information⁴⁶. Furthermore, it is rare that decision situations are neatly structured and ordered such that Commanders can make decisions based only on an assessment, even if additional context is provided with that assessment⁴⁷.

C.2. Operations Assessment Best Practices.

There have been recent attempts to distill best practices for operations assessment. The following describes two such attempts. First, a NATO study produced a guide with a wide set of best practices for dealing with these challenges that can be applied to any assessment. Its purpose is to complement, not replace, the more detailed planning or instructional documents contained in official instructions and field manuals. Instead, the guide attempts to close the gap between the ideal and the reality of assessment by providing insights into the “philosophy” of assessment, highlighting the challenges, and sharing best practices from the field. Second, a recent Doctrinal publication from the United Kingdom described some simple principles for operations assessment.

⁴⁵ Stone, D. (2002). *Policy Paradox: The Art of Political Decision Making*. New York: Norton Press

⁴⁶ Weiss, C. (1979). The many meanings of research utilization. *Public Administration Review* 39(5):426-431.

⁴⁷ Weiss, C. (1980). Knowledge creep and decision accretion. *Science Communication* 1(3):381-404.

C.2.1. Best Practices Guide for Conducting Operation Assessments for Counter-Insurgencies

This section contains a brief description of the 19 articles⁴⁸ in the guide. The articles in the guide are broadly categorized as either assessment tenets (Part One) or assessment methods (Part Two). The assessment tenets seek to clarify the assessment's purpose and objectives. By reminding practitioners of how their assessments can be used to influence an overall campaign strategy, it becomes easier to make the right choices between sources and methods. It also helps practitioners understand how to build and communicate an assessment to influence strategic decisions. Part Two is more tactical in nature, with some of the most common assessment pitfalls, reminders of some fundamentals, and suggestions for dealing with obstinate players or intractable obstacles.

Part One: Assessment Tenets

Article one: Remain true to the assessment's objective. The objective of an assessment is to produce insights pertaining to the current situation, and to provide feedback to support decisions by the leadership. This article discusses how key elements of this objective should guide the assessment development process.

Article two: Take a multi-dimensional perspective. This article describes why it is essential to build the assessment by looking at the environment through multiple perspectives that cross lines of operations and time periods. It also highlights some errors that may arise if the assessment lacks a broad perspective.

Article three: Serve as the bodyguard of truth. Assessment teams develop what may, by default, become the only publicly available, official picture of the campaign. Therefore, assessment teams must serve as the bodyguard of truth and never compromise the integrity of their reports. This article outlines nine key practices that help preserve the integrity of assessments.

Article four: Ensure independence and access. Assessment teams need access to a wide array of information and people in order to perform

⁴⁸ Note, the original source has 20 articles. The list has been modified to reflect NATO terminology. <http://smallwarsjournal.com/jml/art/best-practices-guide-for-conducting-assessments-in-counterinsurgencies>

their job properly. This article describes how to secure independence and access through a partnership between the senior sponsor of the assessment team and individual line of operation owners.

Article five: Nurture the Knowledge Development (Intelligence) – Assessment Partnership. The activities related to knowledge development and assessments often seem remarkably similar, thus generating the potential for confusion or duplication of effort. This article discusses how to build a mutually supporting relationship between the two activities.

Part Two: Assessment Methods

Article six: Establish a terms of reference document. Unclear terms generate confusion in the design of the assessment framework, the analysis of data, and the reporting of insights. Thus, it is in the team's best interests to develop a Terms of Reference document that includes an assessment design and a data collection strategy as soon as possible.

Article seven: Build the assessment framework iteratively, incrementally, and interactively. The assessment framework should be built in stages through a collaborative process to minimize complexity, allow for effective learning, and retain clearly established priorities. It also allows the assessment team to refine the focus and scope of the assessment framework based on lessons learned.

Article eight: Take care over terminology, many use the term “indicator” and “metric” interchangeably, but they may have different meanings.

Article nine: Beware of manipulated metrics. Some metrics can be manipulated by the subjects under observation to send misleading signals to observers, rather than reflecting the reality of the current conditions. This is a particularly high risk for metrics used to promote or demote, or directly redistribute resources. This article contains several examples and suggests ways to detect and minimize such distortions.

Article ten: Develop a manageable set of metrics. There are hundreds of metrics available at any point in time. Thus, it is necessary to establish rules to help select the metrics contributing the most to the operation. This article discusses several screening filters that help practitioners develop a manageable and effective set of metrics.

Article eleven: Retain balance in both metrics and method. Debates on the merits of the narrative report versus summary graphics, the organisational level at which assessments should be performed, and how to preserve the front-line Commander's views within higher level summary assessment products persist in the assessment world. This article describes how to achieve a balanced blend of each alternative that captures the best features of each.

Article twelve: Use field assessment teams. In order to provide actionable information to the decision-maker, assessment insights must be relevant and credible. For critical issues, the only way to achieve this standard is to directly interview front-line units in the field. This article offers an approach that augments the traditional process with the use of field assessment teams dispatched from appropriate levels.

Article thirteen: Use eclectic marginal analysis to bound estimates. When a desired metric is difficult to measure directly, it might be possible to measure the factors that drive the value of the same metric. Under such conditions, marginal analysis can be used with an eclectic set of related metrics to generate a reasonable estimate of the target metric. This section explains the technique and provides some examples of marginal analysis.

Article fourteen: Anchor subjectivity. A degree of subjectivity in assessments is unavoidable. This article discusses methods to minimize the degree of subjectivity, make that subjectivity transparent, and maintain consistency in the way subjective assessments are captured.

Article fifteen: Share data. Every coalition effort faces information sharing challenges. This article discusses important reasons for sharing information and offers some guidelines that promote effective sharing.

Article sixteen: Include host nation data. This article addresses the challenges of using host Nation data and ways to work around the challenges, including host nation data collection systems and the ability of assessment teams to interact with this system.

Article seventeen: Develop metric thresholds properly. This article discusses key guidelines for developing metrics thresholds, including when to adjust levels, how to develop clear definitions of the thresholds, and how to determine when observances of metrics represent a significant change in underlying conditions.

Article eighteen: Avoid substituting anecdotes for analysis. Anecdotes are a useful component of assessments when used properly. Unfortunately, they are often used as substitutes for a solid assessment. The best rule to keep in mind when using anecdotes is that they are generally the starting point for analysis, not the closing argument of an assessment.

Article nineteen: Use survey data effectively. Questions about motivation, satisfaction, degrees of trust or fear, as well as intentions regarding future actions are difficult to measure by monitoring physical activities. Often, this information must be captured by interviews or broader surveys. This article addresses how to manage some of the major concerns associated with using survey data in assessments.

C.2.2. Key Messages and Top Tips

This section is adapted from the UK's Joint Doctrinal Note 2/12: Assessment, dated February 2012⁴⁹.

- Operations assessment should be integral to all operations and from the outset embedded in campaign planning and design at the highest level.
- The assessment is core to effective integrated campaign planning and to increased knowledge.
- Operations assessment is a tool to regularly review, adjust plans and support decision making.
- Understand the situation from the target population's perspective and frame the assessment in their terms.
- Operations assessment that is not based on a top-down information requirement can provide a distorted picture.
- Operations assessments enable Commanders to test assumptions, judge progress, learn and adapt.
- Operations assessment frameworks must be consistently adopted by succeeding leadership and Commanders.
- Military should adopt a comprehensive approach to sharing/exchanging assessment with external actors.

⁴⁹ <https://www.gov.uk/government/publications/joint-doctrine-note-2-12-assessment>

- Operations assessment must recognise the perspectives and influence of the interagency and multinational involvement.
- Assessments must balance the need to maintain a consistent set of metrics to ensure that longer-term effects and trends can be identified, and the need to adapt in order to remain useful in a dynamic environment. It is better to evolve an assessment regime than make wholesale changes to it or to innovate new schemes.
- The process of determining how to measure an effect or objectives enhances conceptual understanding and leads to better designed plans and more insightful objectives.
- Measures of Performance (MOP) give an indication of the extent of activity in execution of the plan.
- Measures of Effectiveness (MOE) give an indication of the success of the plan and results achieved.
- Do not underestimate the resources needed for operations assessment and be prepared for growth as the operation develops.
- Focus on results (the effects), rather than on activity (for example, the number of people trained).
- Assess the impact of civilian and other military activities, including those of host nations and allies.
- Use a balance of quantitative and qualitative input to assessment to inform, support and challenge the Commander's assessment and decision making.
- Regularly review the validity of the assessment in order to understand changes in the operating environment and identify potential unintended consequences.
- To avoid the risk of bias, use information from a wide range of sources.
- Prioritise data collection requirements to both prevent information-overload and minimize risk and cost
- Aim to capture unintended consequences, 2nd and 3rd order effects, as well as interdependencies between different strands of the campaign.
- Assessment planning and execution must be incorporated into formation-level collective training, mission rehearsal

exercises and pre-deployment training, and the lessons-learned process.

Annex D. Examples of Metrics

D.1. Examples of Metrics

D.1.1. Example Metric 1

Plan Elements	Element Values
Effect:	The security situation in country X is restored to the level of 2005 within 12 months.
Criteria:	The security situation refers to the acceptance of the Government as the legitimate authority by the general population.
MOE:	Number of anti-government incidents.
Acceptable Condition:	Average 3 incidents per week over the last 3 months
DROC:	Baseline situation to Acceptable Condition over 12 months as a linear trend

As written, there is no distinction between peaceful demonstrations, violent riots, terrorist attacks or other incidents. Likewise, there is no indicator for scale of the incident. As written, this metric would cover anything from an individual handing out a leaflet to 10,000 rioters burning down buildings. More developed MOEs that describe the situation in more detail could be:

Plan Elements	Element Values
MOE:	The number of violent anti-government incidents
Criteria:	A violent incident is one in which people are injured and require hospital treatment, and/or in which property is destroyed, carried out by members of the public singly or in loose co-ordination, as opposed to individuals who operate as an organised group with a recognised identity.
MOE:	The number of terrorist incidents

Criteria:	A terrorist incident is one that is carried out by an organisation included on the “Operation X List of Terrorist Organisations”
MOE:	The number of peaceful anti-government incidents
Criteria:	A peaceful incident is one in which there are no injuries requiring hospital treatment as a result of direct action either by civilian participants or security forces, and no destruction of property.
MOE:	Number of country X citizens living outside country X with UN refugee status
MOE:	Number of Internally Displaced Persons

This expanded range of metrics can give a much broader overview into the overall security situation and the impact it is having on the general population. However, it will require careful interpretation of the evidence during the assessment process in order to come to an overall conclusion on the status of effect creation.

D.1.2. Example Metric 2

Plan Elements	Element Values
Effect:	The success rate of pirate attacks is reduced below 10% through the self defence capability of merchant ships by the end of 2011.
MOE:	Proportion of attacks on merchant ships that result in a successfully pirated vessel
Desired Rate of Change:	Baseline value on G-Day to Acceptable Condition in Dec 2011 as a linear trend
Acceptable Condition:	10% or fewer attacks result in successfully pirated vessel

Criteria:	<p>A successful pirate attack is one that results in the pirates having control of both the crew (i.e. a hostage situation) and the vessel</p> <p>An unsuccessful attack is one where weapons are discharged by pirates or pirate paraphernalia makes contact with the vessel or pirates get on board the vessel, but the activity is abandoned or disrupted before pirates gain control of the crew and vessel.</p>
-----------	--

This effect focuses on self-defence because this provides the long term strategy that will allow international military forces to withdraw from the region. The contributions made by military forces in parallel to the development of this effect will be addressed by other effects within the plan. While the MOE described here gives the high level overview of effect achievement, it does not assist in understanding the specific self defence mechanisms employed by merchant vessels, and how effective each of these mechanisms are. That may require more detailed analysis and more specific MOE, that examine the importance of factors such as merchant ship size, speed, operating area, weather conditions, time of day etc., assuming that this level of detailed data is available, and whilst ensuring that this more detailed categorisation of the data does not result in sample sizes that are too small from which to draw reliable conclusions.

Plan Elements	Element Values
Action:	Regularly review the adequacy of Best Practice Advice for Merchant Vessels on Self Defence against Pirate Attacks
Criteria:	Best Practice Advice is created through the analysis of the post-incident reports from merchant ships that have been attacked by pirates
MOP:	Proportion of attacks in the last 6 months where correct application of Best Practice Advice would still have been ineffective against the pirates

Acceptable Condition:	7% or fewer attacks where correct application of Best Practice Advice would still have been ineffective.
Desired Rate of Change:	Baseline situation to Acceptable Condition within 18 months along a trend that steps down in 6 month intervals.

This metric is an example of quantifying subjective data, where the underlying assessment about the effectiveness of Best Practice Advice for each individual incident is likely to be based on the judgement and experience of the merchant ship master combined with specific questioning from the actor responsible for collating the information. Combining the results of individual incident assessment gives the quantitative result.

The Acceptable condition would be derived from a combination of other elements of counter-piracy operations, based on what is likely to be an acceptable situation based on assumptions of likely future pirate behaviour and without the availability of security forces in the region. The DROC is based on the fact that it will take time to gather and analyse the data from individual incidents, and publication of updated Best Practice Advice is likely to be an occasional occurrence. Occasional publication of updated advice combined with the fact that it would take time for merchant ships to digest and update their own capabilities is recognised in the fact that the DROC changes in steps rather than gradually.

D.1.3. Example Metric 3

Plan Elements	Element Values
Effect:	The Air Force of country X is incapable of disrupting the flow of friendly forces into the Sea Ports of Disembarkation (SPODs) of country Y.
MOE:	Number of friendly force troops transited through SPODs of country Y
MOE:	Number of Linear Metres of friendly force equipment transited through SPOD of country Y

Desired Rate of Change:	Force flow as defined in OPLAN
Acceptable condition:	Final force level delivered by sea as defined in OPLAN
Supporting Metrics:	
MOP:	Number of sorties flown by country X air force over the Ocean to within 250 nautical miles of the country Y coastline
MOP:	Number of sorties flown by country X air force that include an incident involving shipping deploying friendly forces to country Y
Criteria:	An incident is an event that results in shipping deploying friendly forces changing speed and/or course such that its planned arrival time at the country Y SPOD is delayed by at least 12 hours

This effect is based on the initial planning assumption that the country X Air Force poses a credible threat to shipping. The MOEs associated with this effect use the rate of friendly force flow as the primary indicator, implying that if force flow is acceptable, then by inference the effect has been achieved. The consequence of initially having only this high level metric is that, should it show that there is a problem with friendly force flow, then more detailed analysis is required to determine whether the country X Air Force is the primary cause of that disruption. Also, by implication, there are activities defined elsewhere in the plan that describe the action that friendly forces must take in order to render the country X Air Force ineffective at least in terms of the threat to shipping used by friendly forces.

Plan Elements	Element Values
Action:	Maintain continuous Combat Air Patrol (CAP) at points Echo and Foxtrot while friendly forces deploy through the SPODs of country Y

MOP:	Proportion of time for which CAP is in place at point Echo Proportion of time for which CAP is in place at point Foxtrot
Acceptable Condition:	CAP is in place for at least 90% of the time at point Echo CAP is in place for at least 90% of the time at point Foxtrot

This is a straightforward action, which is either being performed or not at any particular time. The Acceptable Condition being defined as less than the perfect solution reflects the realities of factors such as aircraft availability and serviceability, weather conditions, etc., and can reflect the Commander's judgement about what is an acceptable level of risk to the mission. As the intention is that these actions are conducted continuously over the period of interest, there is no DROC defined, as there should be no intended gradual change in the situation.

D.1.4. Example Metric 4

Plan Elements	Element Values
Effect:	Local economies recover to pre-crisis levels within 24 months.
Criteria:	Local economies refer to the economies of individual provinces within country Alpha, as IO/NGOs operating within country Alpha have action plans that treat provinces individually
Overarching MOE:	As defined in the United Nations Mission In Alpha (UNMIA) Country Development Plan
Acceptable Condition:	As defined in the UNMIA Country Development Plan
Desired Rate of Change:	As defined in the UNMIA Country Development Plan
Military MOE:	<p>Individual criteria for each province based on the 3 most important aspects of the provincial economy</p> <p>Coastal Province: Value of trade at Town A fish market Value of international goods imported via Town A port Value of Country Alpha produced goods exported through Town A port</p> <p>Capital Province: Number of individuals employed as Civil Servants in Capital Number of flights by international airlines to Capital airport Hotel occupancy rate in Capital</p> <p>Inland Province: Number of hectares of arable land in use Value of trade at Town B vegetable market Number of tourists visiting Inland Province Wildlife Reserve</p>

While the state of the economy in the nation of interest in this example is an important factor in determining the overall national situation, and therefore may impact on military decisions directly (is a failure in the recovery of local economies due to a detrimental security situation) and indirectly (is the country stable enough that international military forces can withdraw), the main responsibility for creation of the effect lies with non-military actors. While the mechanisms by which the effect is created may be beyond direct military control, and therefore factors such as the scope of the assessment (i.e. by Province), the Acceptable Condition and Desired Rate of Change will be dependent on the plans and intentions of those non-military actors, it is still possible for the military to take an independent view on whether they believe the effect is actually being created through the use of their own MOE.

Plan Elements	Element Values
Action:	Collect data on the value of Trade at Town A fish market.
MOP:	Proportion of days for which data is collected
Acceptable Condition:	Data is collected on 30% of days.

This example shows a situation where an action may need to be undertaken specifically to support the assessment process, in gathering data that would not otherwise be available. A relatively low level is set for the Acceptable Condition on the assumption that any change in the data being collected is slow compared to the periodicity of the data collection, although care must be taken to ensure that any specific local bias can be considered (e.g. it is traditional to eat fish on a Friday).

D.1.5. Example Metric 5

Plan Elements	Element Values
Effect:	Rebel forces are dissuaded from conducting operations within Capital whilst National Army capability is rebuilt to pre-crisis levels.
Criteria:	<p>Rebel forces refer to any armed group who are not National Government or International Military Security forces. Rebel operations are individual identifiable actions designed to further the cause of any non-government group, and includes both political actions designed to destabilise the government (e.g. attacks on government facilities, intimidation of civilians attempting to use government services), and criminal actions designed to support Rebel capabilities (e.g. theft in order to obtain funds for weapon purchases).</p> <p>National Army capability refers to the number of soldiers who are available for active duty (i.e. trained, issued with personal equipment and weapon, not on leave or medically unfit etc), the number of pieces of heavy equipment (i.e. transport, heavy weapons etc.), plus infrastructure (i.e. barracks, training areas etc.)</p>
MOE:	Number of operations conducted within Capital by Rebel forces.
Acceptable Condition:	Zero operations by Rebel forces.

The effect focuses on the deterrence of Rebel forces rather than the development of National Army capability. Assessment of this effect can be relatively simple, as the defined metric only provides a high level overview of Rebel activity. More detailed metrics would need to be defined in order to distinguish between criminal and political acts if required to inform decision making in other areas of the operation. While the criteria for National Army capability is defined here in order

to fully understand the effect, the development of that National Army capability would be covered by other effects, actions and associated metrics. The only consideration associated with National Army capability development is that it defines the time boundaries applicable in this situation. There is no Desired Rate of Change associated with this metric, because the effect is a condition that must be achieved immediately at the start and then maintained throughout the duration of the period during which National Army capability is being built, rather than one that needs to be attained at some point during this phase of the operation.

Plan Elements	Element Values
Action:	Maintain security checkpoints at the North, South, East and West major road entrances to Capital
MOP:	Proportion of time for which all Capital checkpoints are manned Proportion of time for which checkpoint North is manned. Proportion of time for which checkpoint South is manned. Proportion of time for which checkpoint East is manned. Proportion of time for which checkpoint West is manned.
Acceptable Condition:	All checkpoints must be manned at least 80% of time. Checkpoint North must be manned 95% of the time Checkpoint South must be manned 95% of the time Checkpoint East must be manned 95% of the time Checkpoint West must be manned 95% of the time

A combination of overall and individual MOPs are used in this situation, as the overall MOP will indicate the proportion of time for which Capital is vulnerable, assuming Rebel forces have the capability to monitor checkpoints and communicate which and when are unmanned, while the individual checkpoint MOPs will give more

specific information on the performance at each. The Acceptable Condition being defined as less than the perfect solution reflects the realities of factors such as manpower availability, priority of other actions etc., and can reflect the Commander's judgement about what is an acceptable level of risk to the mission. As the intention is that these actions are conducted continuously over the period of interest, there is no DROC defined, as there should be no intended gradual change in the situation. A high level of granularity in the MOP may be required to determine patterns of behaviour that may present specific vulnerabilities (e.g. a specific checkpoint is always unmanned late at night).

Annex E. Polling to Support Operations Assessment

Polling can provide an understanding of public perceptions and attitudes, and can therefore, when conducted over a period of time, show how opinions have or have not changed. This annex provides guidance on using polling data in support of operations assessment. It describes the circumstances under which it is appropriate to place confidence in data collected by contractors who undertake polling in a theatre of operations. Results of polling can be used at tactical, operational and military strategic levels.

Polling has been used in recent operations assessments in Afghanistan with relative success; however, scepticism of the findings from some polls prevailed. This was a result of misunderstanding of the polling methods used. In addition, in some areas there was an over reliance on polling data to support operations assessments, ignoring or discarding the caveats of the poll. The lessons learned from previous experiences are important and reflected within this annex to avoid the common pitfalls.

Conducting a poll for the purposes of operations assessment involves a number of stages. Figure E-1 presents an overview of the stages and workflow of a polling process demonstrating the relevant parties involved in each stage. The number of interested parties and their responsibilities depend on the scope and type of poll. Parties involved can include the assessment staff, sponsors (including who holds the cash –J8), stakeholders (who are the end user of the polling products), subject matter experts (those qualified to support the development of a questionnaire) and polling contractor(s) (the organisation who will field and analyse the poll). .

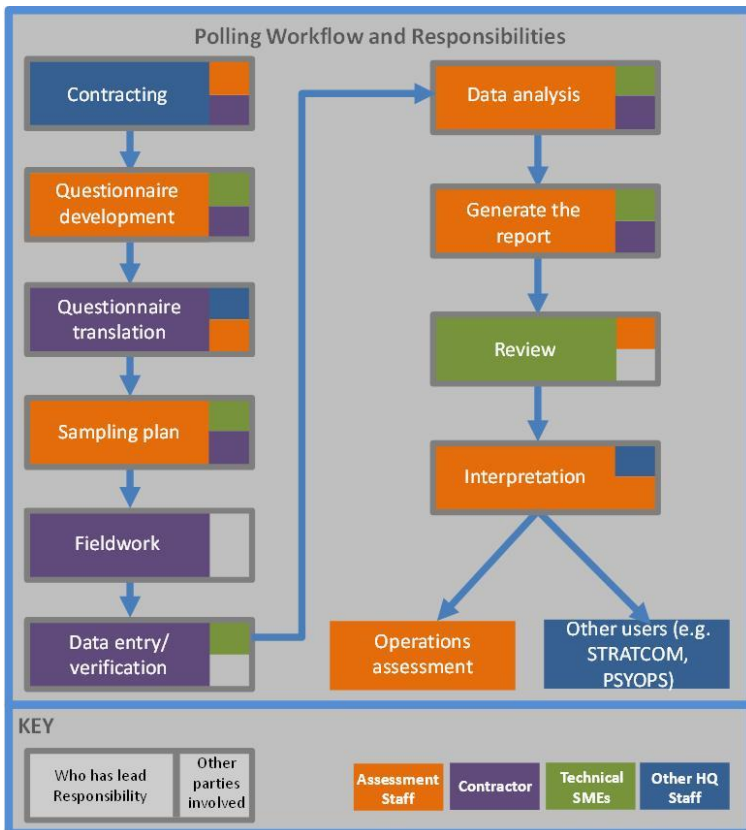


Figure E-1: Polling workflow and Responsibilities

Conducting a poll to a high standard in impermissible environments is a challenge, particularly when measuring the opinions of people from a different culture interacting with NATO military presence. Polling contractors need to do the best piece of research, under the circumstance. As such, reported findings should always be interpreted within the context of any methodological concerns or flaws, and details regarding the methods used and problems encountered should always be readily available.

Summarised below are the key issues and questions that should be asked when using polling to support operations assessment. Use this information as a guide to help assess the quality of a polling activity

and determine as a consequence whether findings should be used to assess the impact of a NATO mission.

E.1. Polling Initiation, Management and Oversight

Assess the suitability of opinion polling. Assess whether opinion polling is the right data collection method. Public opinion polling will not provide data that gives a deep understanding of public sentiment and its underlying causes.

Engage subject matter experts early. As early as possible in the design of a poll, opinion polling experts and cultural experts should be consulted. Experts may be found elsewhere within a HQ, within other NATO organisations, NATO nations' government departments, NGOs, IOs or academia.

Identify previous polls conducted prior to NATO deployment. Do your research and identify what polls have been conducted within a country. Not only can this provide insight to the potential methodological issues that NATO could face, but also give a baseline assessment to support the assessment design in the planning phase.

Be aware of the host nation's policy regarding public opinion research. Some nations impose certain legislation regarding public opinion research. When possible, consultation with the Host Nation government should always be sought at the initiation of any NATO poll.

Select a trusted contractor. Request proof of professional qualifications, reference, evidence of prior related work and membership in relevant professional associations can provide some assurance in the selection of a trusted contractor. Contracts should include options for early termination of the contract where necessary.

Create a working group that includes stakeholders and subject matter experts. In preparation for an initial poll, all relevant stakeholders and technical advisors should be identified. They should be consulted regarding poll design and informed of the poll findings. As the polling campaign evolves, the group should meet regularly to review the instrument, sampling plan and results to recommend adjustments. The group may include military planning and assessment staff, information operations staff, representatives from other relevant government departments and operational analysts.

Select a flexible initial design. Expect to learn lessons from early polls, that may lead to changes in the instrument, survey design or potentially even the choice of the contractor providing field teams. As such, the initial contract should cover a single wave of polling, or have flexible termination clauses. It should also allow for changes in the instrument and sampling plan.

Deliver an initial poll. A first poll can serve as a trial before a commitment is made to a long-term polling campaign. Such a poll can be used as a proof of concept to demonstrate the utility of polling to potential sponsors.

Consider conducting a poll before D-Day. In order to develop baseline assessments, and where permissible, it may be possible to establish a poll before D-Day. Polls may be fielded out of country by telephone or internet, where permissible.

Consider reachback to permanent HQ to manage the poll program in order to maintain program continuity. Deployed HQs can suffer from frequent turnover, which impacts continuity in survey design and analysis.

E.1.1. Have You Polled Enough People?

An adequate sample size is needed depending on the level of analysis required.

The table demonstrates the required sample sizes for populations between 1000 to 1000000, for a 5% margin of error at 95% confidence level. NB. These errors assume the sample is *fully* representative of the population from which it is drawn!

Population	Sample size
1000	278
10000	370
50 000	381
100 000	383
250 000	384
500 000	384
1 000 000	384

QUESTIONS TO ASK

What is the sample size? Is it adequate for the analysis required?
 What is the 'non response' rate? I.e. how many people refused to respond?
 Is the sample likely to be representative of the nation/province/district?
 What issues might have affected data (for example, people living in particular areas that could not be reached to poll)?

E.1.2. Whose Opinions Are You Examining?

Random sampling is necessary to reduce bias in data, and stratified sampling methods ensure that the views of relevant sub populations are represented. A Kish grid⁵⁰ (or similar tool) should be used to select a household and a person within this to take part.

QUESTIONS TO ASK

What sampling methods were used? Was it stratified random sampling?

Were sampling methods clearly described by those responsible for the poll?

E.1.3. Difficulties In Accessing Participants

Within a conflict environment, people located in certain regions may be selected to participate via the sampling method, but cannot be polled due to reasons such as remoteness, natural hazards or security issues. It is important the polling contractor report when such problems arise; this information is also a good indicator for assessment purposes.

QUESTIONS TO ASK

Did the polling company report such issues? If so, how did they overcome these?

Were findings reported in context? That is, were methodological flaws and limitations such as flooding, remoteness or security reported?

E.1.4. Data Collection

For trend analysis, data needs to be gathered at close enough intervals to be able to explore reasons for any changes observed. Questionnaire items also need to stay consistent over time- if they do not, trend analysis should never be conducted on these items. New questions can be added, but consistency is essential for quality analysis.

QUESTIONS TO ASK

How often were polls conducted? Were they at regular intervals, e.g. quarterly or monthly?

Did the questions asked remain consistent over time?

⁵⁰ A Kish grid is a tool to ensure that all individuals in a household have an equal chance of selection for a survey (consult a specialist analyst for more information)

E.1.5. Questionnaire Construction

The questionnaire needs to comprise carefully worded and culturally appropriate questions. In this respect advice from a social scientist should be sought. Response options need to cover a range of choices including 'Don't Know' and 'Refuse to answer'.

QUESTIONS TO ASK

Was the actual questionnaire used available with results?

Was it easy to use, well worded, and did it avoid leading and/or double barrelled questions? What response options were available?

Were 'neutral', 'don't know', neither agree nor disagree' options provided when appropriate?

Here are some tips on best practice for question construction.

Avoid the use of confounding or double barrelled questions. For example, responses to 'How helpful and appropriate do you find the Police?' may vary because 'helpful' and 'appropriate' are two different concepts requiring two separate questions.

Avoid leading questions. For example, by starting a question with: 'Do you agree that...', this may encourage respondents to tend to agree.

Include neutral options. This ensures data is collected accurately and reflects peoples' opinions, rather than being forced into a response that may not reflect their view point. For example 'Neither agree nor disagree' or 'Don't know' may be suitable.

Keep questions consistent across polls over time. Wording of questions should remain constant across polls. Changing the wording of questions or deleting questions can have a negative impact on the ability to conduct trend analyses which are critical for operations assessment.

Consider the length of the poll. Polls that are too long will affect the quality of responses. For example, when participants are faced with lots of questions they are likely to lose interest or become frustrated and therefore more likely to respond arbitrarily towards the end.

Consider the order that the questions are presented in. Some earlier questions may have a 'priming' effect on responses to later

questions. It may be pertinent to randomise the order of the questions for different participants.

E.2. Questionnaire Piloting and Testing

The questionnaire should always be piloted with a small sample of the target population, in order to check that questions are understood and culturally appropriate. Checks must also be made to ensure that questions and response options have been translated correctly, particularly by having them translated back by a different linguist.

QUESTIONS TO ASK

Was the questionnaire piloted?

Was it back translated? How did the polling contractor cope with different dialects?

E.2.1. How Should Participants Be Interviewed?

Polling staff should be well trained in how to administer the questionnaire. It is important that they know how to develop rapport with participants in order to reduce the chance of socially desirable responses being given. It is also important that they are trained to emphasise the participants that there is no 'right' or 'wrong' answer.

QUESTIONS TO ASK

Were staff trained in interviewing skills and how to administer the questionnaire?

Were they likely to understand the importance of rapport building and were they trained to do so?

Was time allowed for this phase of the interview?

E.2.2. Are the Questions Culturally Sensitive?

Cultural norms will influence how polling activity is perceived by participants. It needs to be perceived as credible to elicit valid and reliable data.

QUESTIONS TO ASK

Was a cultural advisor consulted?

Were the staff who conduct the polls trained in cultural norms? Are the staff locally employed?

Was the polling conducted under local cultural norms?

E.3. Quality Controls

No matter how much training interviewers might receive, there is no guarantee that they will administer the sampling methods and/or questionnaire correctly, and in some cases data may even be fabricated. Therefore, the polling contractor needs to implement checks such as data entry checks (getting two independent people to input data then checking to ensure that these are the same) scanning data for patterns that indicate fabrications, and back checking (when a supervisor/member of staff calls – in person or by telephone – a household, and ensures that the participant who was reported to have been interviewed had actually been interviewed).

QUESTIONS TO ASK

Were data entry checks made? Was the data examined for possible fabrication?

Did the organisation employ back checks on reported interviews?

E.4. Data Analysis

The credibility and value of a polling program rests on the quality of the gathered data. Continuous verification and application of quality controls to the data are important. Subject Matter Experts can inspect and verify the data using proven statistical techniques.

It is essential to use appropriate software tools to handle polling data. When possible and practical, the use of standard statistical software such as the NATO accredited tool SPSS (Statistics Package for Social Scientists) is recommended. If this is unavailable, open source tools like R (R-project) is recommended for use on a standalone computer. All data and results should be always made available in standard formats (.csv or .xls) in order to facilitate data sharing with partners who may not have access to statistical software. Although often labour intensive, invest in resources to combine datasets from multiple waves into a single data set to facilitate longitudinal analysis.

E.5. Interpreting Results

Be aware that rather than reflecting a change in opinion, unexpected findings may be a result of the question being misunderstood (by the interviewer and/or interviewee) misinterpreted responses, cultural factors (e.g. reluctance to criticise authority) and or an assumption that

the interviewer is connected with the government (therefore responses may not be over critical). It is stressed that there is a need to verify and validate the findings from polling with other data sources such as atmospheric or media reporting to enhance the quality of campaign assessment.

QUESTIONS TO ASK

Are findings supported by other sources, organisations and so on?

If not, what alternative reasons for unusual findings might be made? For example, might the methodology have affected the data in some way?

Was a local, independent and impartial contractor employed? This can minimise potential bias.

E.6. Presenting Results

The impact of polls in any operational HQ is dependent on the effectiveness of the presentation of the results. The following provides key tips in ensuring high impact of polling results:

Deliver timely results. Polling results should be reported quickly in order that they remain relevant. Setting up a formal structure for review will assist in ensuring timeliness.

Emphasise trends when reporting results. Polling results from conflict zones are likely to be subject to systematic biases. However, trends over time or regional variations are less sensitive to these biases if they are assumed to be unchanged in either space or time. Therefore, trends should be emphasised in the reporting of polling results.

Provide a synthesis of the numerical results. Spell out the results in text. Providing only a list of numerical results is highly undesirable. Users may not have the inclination to comb through data presented in such a way. More importantly the user may be subject to confirmation bias (i.e. the propensity to favour results that confirms their previously held beliefs).

Present results succinctly. Results should be presented in a succinct and intuitive way, avoiding unnecessary statistical jargon.

Be careful drawing causal inference. Caution should be used when interpreting trends in the data in light of events in the theatre of

operations and particularly inference with respect to operations assessment.

Report potential sources of error and biases. Errors and potential biases should be clearly described in reporting the results.

E.7. Triangulation

Triangulation is the process of using more than two methods in order to examine/double check results. Triangulation helps to overcome problems stemming from reliance upon a single logic, method, source, or set of data.

It is important to ensure that a data collection plan is drawn in coherence with the design of any polling activity from the outset.

If the results from the different methods and sources converge, inference about the measures of effect and the magnitude of the effect will be stronger.

E.8. Summary

Polling can provide an excellent source of situational awareness that can be used for operations assessment. However this assertion comes with a serious caveat: The methods selected to field a poll must be well understood in order to determine the quality of any such data. Conducting any kind of poll to a high standard is not an easy task; it is even more difficult when the aim is to measure the opinions of people from a different culture within a conflict environment.

A key take home message is that whilst methodological rigour is important, collecting data in a military environment is inevitably going to be difficult. Some trade-off between methodological rigour and the need to understand local opinion – is going to be necessary. Polling contractors need to conduct the best poll they can, *under the circumstances in which they work*. As such, findings should be interpreted within the context of any methodological concerns or flaws, and details regarding methods used and problems encountered should always be readily available.

Annex F. Guidance for Annex OO in the Operational Plan

F.1. Introduction

The success of operations assessment will be predicated upon the clear and concise orders set out in the operational plan prior to execution of an operation. Annex OO to the operational plan is reserved for the use of operations assessment (See NATO Comprehensive Operations Planning Directive (COPD)). This annex provides general guidance on the information that should be published in any given Annex OO.

F.2. Annex OO Template

The format of Annex OO should follow the guidance as given in the COPD, using the NATO standard six paragraph format: Situation, Mission, Concept of Operations, Execution, Service & Support and Command & Signal. The following template serves as a handrail for staff officers to ensure an effective operations assessment annex to an OPLAN, OPORD or CONPLAN. It provides suggested headings and recommended information for inclusion. As a minimum all headings in the annex should be published at the same time as the main body of the plan (It is likely that the assessment plan will expand and refine over time and should be updated through the Fragmentary Order process accordingly).

ANNEX OO – OPERATIONS ASSESSMENT

1. SITUATION

a. General. Introduction to operations assessment, its purpose within the HQ, relationship to the plan and the key references used in the design of the assessment plan.

b. Purpose. The purpose of the annex.

2. MISSION. A clear, concise statement which states the operations assessment mission, with a clear purpose in support of Commander's decision making.

3. CONCEPT OF OPERATIONS

a. General Concept for Operations Assessment. The general overview of the assessment will be conducted

including the MOEs/MOPs, data collection, how the data will be analysed to develop outputs, where the assessments will be used and what decisions the assessments are likely to support. Include reference to how lessons learned will be captured⁵¹ and the assessment refined.

b. Operations Assessment Model/Process. A schematic drawing representing an overview of the process of operations assessment within the command.

c. Operations Assessment Results. How will the assessment products be presented? Where and who will use the output from the assessments?

d. Data Collection Plan. Reference to how data will be collected using the data collection plan detailed in Appendix I.

4. EXECUTION

a. Operations Assessment Battle Rhythm. How the operations assessment will be executed with a battle rhythm and its relationship with the wider HQ battle rhythm.

b. Co-Ordinating Instructions

i. Subordinate Command Actions. Actions or responsibilities for subordinate Commands.

ii. Supporting Command Actions. Actions or responsibilities for supporting Commands.

iii. Host Nation Requests. Requests to the Host Nation for support. Identify overlaps with Host Nation assessment capabilities.

iv. Civilian Organisations Requests. Requests to Civilian Organisations for support. Identify overlaps with Civilian Organisations assessment capabilities.

c. Use of TOPFAS or other Operations Assessment related software. How the assessment will be executed using

⁵¹ See Bi-SC Directive 80-006 Lessons Learned for guidance on implementing an “assessment” local lessons learned process

software applications, including databases and tools such as the Campaign Assessment Tool within TOPFAS.

5. SERVICE SUPPORT

a. Finance. If any service contracts are to be established related to operations Assessment, for example polling, detail plans for contracting here.

6. COMMAND & SIGNAL

a. Command & Control. Describe the relationship with other assessment cells.

b. Liaison & Co-ordination. Describe how to deal with issues and who the key Points of Contact are within the Command.

c. Reporting. Detail key reports and timings for submission.

SIGNATURE BLOCK

APPENDIX LIST

APPENDIX I – DATA COLLECTION PLAN

Annex OO writers should be careful when writing very specific MOEs and MOPs because they may need to change after the approval of the OPLAN. As the OPLAN reviews take time, the annex may become obsolete in some essential aspects, requiring adjustment through mechanisms other than OPLAN review. Bearing this in mind, a plan which includes the following information for the purposes of data collection:

MOE/MOP with associated planning elements such as Operational Objective, Decisive Condition, Supporting Effect, task etc. Include all reference numbers.

Detailed description of MOE/MOP including definitions.

Goals of the MOE/MOP (including AC, DROC)

Type of data being collected (including units of measurement)

Data source

Officer of Primary Responsibility for data collection

Data format to be reported in

Frequency data to be reported

Annex G. Operations Assessment – References

This list of references serves to provide the assessment staff officer with additional texts that supplement this handbook. These texts offer different perspectives and views on the art of operations assessment.

“Assessing Counterinsurgency and Stabilization Missions”, Jason Campbell, Michael O’Hanlon, Jeremy Shapiro, Brookings Institute, 2009,

http://www.brookings.edu/~media/research/files/papers/2009/5/counterinsurgency%20ohanlon/05_counterinsurgency_ohanlon.pdf

“Assessing Progress in Military Operations: Recommendations for Improvement”, produced by United States Joint Forces Command for Multinational Experiment 6.

“Embracing the Fog of War: Assessment and metrics in counterinsurgency”, Ben Connable, RAND, 2012,

<http://www.rand.org/pubs/monographs/MG1086.html>

“Guidance on Evaluating Conflict Prevention and Peace building activities”, OECD/DAC,

<http://www.oecd.org/dac/evaluation/dcdndep/39774573.pdf>

“In the Balance: Measuring Progress in Afghanistan”, Courtney, M., Riddell, H. Ewers, J., Linder, R. and Cohen, C. CSIS, 2005,

<http://csis.org/files/media/csis/pubs/inthebalance.pdf>

“UK JDN Operations Assessment: Monitoring and Evaluation in a Stabilisation context”,

<https://www.gov.uk/government/publications/joint-doctrine-note-2-12-assessment>

“Operations Assessment in Afghanistan is Broken: What is to be done?”, Stephen Downes-Martin, Naval War College Review, 2011.

<http://www.usnwc.edu/getattachment/e8262311-4922-4b8c-b14c-8714992130ba/Operations-Assessment-in-Afghanistan-Is-Broken--Wh.aspx>

“Measures of Effectiveness for the Information-Age Army”, Richard E Darilek, Walter L Perry, Jerome Bracken, John Gordon IV, Brian Nichiporuk, RAND, 2001,

http://www.rand.org/pubs/monograph_reports/MR1155.html

“Measuring and Managing for Results in Fragile and Conflict-Affected States and Situations”, UK Stabilisation Unit Guidance Note, <http://www.stabilisationunit.gov.uk/attachments/article/523/Interim%20Guidance%20Note%20-Measuring%20and%20Managing%20for%20Results%20in%20Conflict-Affected%20and%20Fragile%20States%5B1%5D%5B1%5D.pdf>

“Measuring Effectiveness in Complex Operations: What is Good enough?” Sarah Jane Meharg, Canadian Defence and Foreign Affairs Institute, 2009, <http://www.cdfai.org/PDF/Measuring%20Effectiveness%20in%20Complex%20Operations.pdf>

“Measuring Effectiveness in Irregular Warfare”, James Clancy, Chuck Crossett, 2007, <http://strategicstudiesinstitute.army.mil/pubs/parameters/Articles/07summer/clancy.pdf>

“Measuring Progress in Conflict Environments (MPICE), US Army, 2010. [http://www.usip.org/files/resources/MPICE_final_complete%20book%20\(2\).pdf](http://www.usip.org/files/resources/MPICE_final_complete%20book%20(2).pdf)

“Military Operational Measures of Effectiveness for Peacekeeping Operations”, Joseph Anderson, Naval War College, 2000, <http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA381728>

“No Sure Victory: Measuring US Army Effectiveness and Progress in the Vietnam War”, Gregory A. Daddis, New York: Oxford University Press, 2011

“Operational Assessment – The Achilles Heel of Effects-Based Operations?”, Christopher W. Bowman, Naval War College, 2002, http://www.au.af.mil/au/awc/awcgate/navy/ebo_bowman.pdf

“Opinion Polling in Support of the Canadian Mission in Kandahar”, Eles, P.T, Vincent, E., Vasiliev, B., Banko, K.M. Defence Research and Development Canada, 2012

“Progress or Peril? Measuring Iraq’s reconstruction” Barton, F. and Crocker, B, CSIS, 2004. http://csis.org/files/media/csis/events/040908_progressperil_executivesum.pdf

“Teaching New Dogs Old Tricks: Can the Hamlet Evaluation System Inform the Search for Metrics in Afghanistan?”, David Gayvert, Small Wars Journal, 2010, www.smallwarsjournal.com/blog/journal/docs-temp/531-gayvert.pdf

“The Uncertain ‘Metrics’ of Afghanistan (and Iraq)”, Anthony Cordesman, Center for Strategic and International Studies, 2007, http://csis.org/files/media/csis/pubs/070521_uncertainmetrics_afghan.pdf

“Understanding Measures of Effectiveness in Counterinsurgency Operations”, Douglas D Jones, 2006, <http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA450589&Location=U2&doc=GetTRDoc.pdf>

Annex H. Acronym List, Glossary

AB	Assessment Board
AC	Acceptable Condition
AROC	Actual Rate of Change
AWG	Assessment Working Group
C2	Command and Control
CAP	Combat Air Patrol
CAT	Campaign Assessment Tool
CCIR	Commander's Critical Intelligence Requirements
CCOMC	Comprehensive Crisis and Operations Management Centre
CIMIC	Civil-Military Cooperation
CJSOR	Combined Joint Statement of Requirements
CMI	Civil Military Interaction
CMPS	Civil Military Planning and Support
CONOPS	Concept of Operations
COPD	Comprehensive Operations Planning Directive
DAC	Development Assistance Committee
DC	Decisive Condition
DROC	Desired Rate of Change
EU	European Union
FRAGO	Fragmentary Order
HQ	Headquarter
HN	Host Nation
IFC	Intelligence Fusion Centre
IKM	Information and Knowledge Management
IMF	International Monetary Fund
IO	International Organisation
ISAF	International Stabilisation for Afghanistan Force
JCO	Joint Coordination Order

JOA	Joint Operation Area
JOPG	Joint Operations Planning Group
KD	Knowledge Development
LEGAD	Legal Advisor
M&E	Monitoring and Evaluation
MOE	Measure of Effectiveness
MSO	Military Strategic Objectives
NAC	North Atlantic Council
NATO	North Atlantic Treaty Organisation
NCIA	NATO Communications and Information Agency
NCMP	NATO Crisis Management Process
NGO	Non-Governmental Organisations
ODA	Official Development Assistance
OGD	Other Government Departments
OECD	Organisation for Economic Cooperation and Development
OPG	Operational Planning Group
OPLAN	Operation Plan
OPORD	Operations Order
OPP	Operations Planning Process
OSCE	Organisation for Security and Cooperation in Europe
PIR	Priority Intelligence Requirements
PMESII	Political, Military, Economic, Social, Information, Infrastructure
PMR	Periodic Mission Review
PROC	Predicted Rate of Change
RBM	Results Based Management
ROC	Rate of Change
SACEUR	Supreme Allied Commander Europe
SHAPE	Supreme Headquarters Allied Powers Europe
SITREP	Situational Report

SME	Subject Matter Expert
SOP	Standard Operating Procedure
SPOD	Sea Port of Disembarkation
SPSS	Statistics Package for Social Scientists
STRATCOM	Strategic Communications
TOPFAS	Tools for Operations Planning Functional Area Services
UN	United Nations

Acceptable Condition (AC)	A defined level for the metric at which a desirable situation has been achieved.
Activity	The process of engaging any instrument at each level in the in the engagement space in order to create (a) specific effect(s) in support of an objective
Actual Rate of Change (AROC)	The Rate of Change measured by the assessment process during the execution of an operation.
Bias	Prejudice in favor of or against one thing, person or group compared with another, usually in a way considered to be unfair
Causality	The relationship between cause and effect
Choropleth	A thematic map in which areas are shaded or patterned in proportion to the measurement of the statistic being displayed on the map (e.g. population density, per-capita income)
Civil-military cooperation	The coordination and cooperation, in support of the mission, between the NATO Commander and civil actors, including the national population and local authorities, as well as international and national, and non-governmental organisations and agencies (Source: AAP-6)
Comprehensive operations assessment	An assessment of progress that includes the entire range of factors in the operational environment, including factors outside the security-domain.

Confidence Interval	A range of values around a statistic (e.g. the mean), with a certain probability of the true value of that statistic.
Correlation	The degree to which two or more measurements show a tendency to vary together
Crisis Management	The coordinated actions taken to defuse crises, prevent their escalation into armed conflict and contain hostilities if they should result. (source: AAP-6)
Data Collection Plan	A plan which describes the attributes of the data to be collected, the data source, how frequently it is collected, the format the data should be recorded and analysed and assigns responsibilities for data collection. The Data Collection Plan is published as an ANNEX to the OPLAN in ANNEX OO.
Decisive Condition	A combination of circumstances, effects, or a specific key event, critical factor, or function that when achieved allows a Commander to gain a marked advantage over an opponent or contribute materially to achieving an operational objective (source: AJP-01(D))
Desired Rate of Change (DROC)	A Rate of Change agreed during planning that indicates the intentions for creating an effect or achieving an objective in a specified period of time, or in application of resources as they become available in order to undertake an action.
Deterministic	Output of a system that can be predicted with 100% certainty.
Effect	A change in the behavioural or physical state of a system (or system elements), that results from one or more actions, or other causes. (source: COPD v2)
End State	The NAC statement of conditions that defines an acceptable concluding situation for NATO's involvement. (source: COPD v2)

Evaluation (Monitoring & Evaluation)	The process of determining merit, worth or value of an activity, policy or program. It consists of the systematic and objective assessment of an on-going or completed project, program or policy, its design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, development efficiency, effectiveness, impact and sustainability.
Goal	Used to judge the achievement of results or activities.
Joint Coordination Order	An order that provides further amplification to the Commander's Direction and Guidance in the OPLAN in order to keep the campaign on track.
Knowledge Development	A proactive, collaborative and iterative process carried out at all levels of NATO, drawing on NATO and non NATO entities, to develop and then maintain a holistic understanding of complex environments in support of NATO political and military authorities' decision making. (source: COPD v2)
Line of Operation	In a campaign or operation, logical line(s) linking effects and decisive conditions in a time to an objective. (source: COPD v2)
Linear Extrapolation	Creating a tangent line at the end of the known data and extending it beyond that limit.
Logical framework (Logframe)	An analytical tool used to plan, monitor, and evaluate (civilian) projects.
Margin of Error	A statistic that expresses the amount of random sampling error in a poll's results. The larger the margin or error, the less confidence one should have in that poll's reported results are close to the true figures (for the whole population). A margin of error is given whenever a population is incompletely sampled.
Measure of Effectiveness (MOE)	A metric used to measure a current system state.

Measure of Performance (MOP)	A metric used to determine the accomplishment of actions.
Metric	A system of measurement – for the purposes of operations assessment this is considered either a MOE or MOP.
Monitoring and Evaluation	Civilian organisations’ equivalent of NATO Operations Assessment.
Monitoring (Monitoring & Evaluation)	A continuous function that uses a systematic collection of data on specified indicators to provide management and primary stakeholders of an intervention with information regarding the use of allocated funds, the extent of progress, the likely achievement of objectives and the obstacles that stand in the way of improved performance.
Objective (in a plan)	A clearly defined and attainable goal to be achieved. (source: COPD v2)
Objective (data and analysis)	Facts and the precise measurement of things or concepts that actually exist.
Operation	A sequence of coordinated actions with a defined purpose. Notes – NATO operations are military, NATO operations contribute to a wider approach including non-military actions. (source: AAP-6)
Operations Assessment	The activity that enables the measurement of progress and results of operations, and the subsequent development of conclusions and recommendations that support decision making.
Operational Analysis/ Operations Research	The application of scientific methods to assist executive decision makers. (source: AAP 6)
Operational Design	The fusion of the operational framework and the Commander’s initial intent.
Poll/Polling	A method for recording the opinion of a population.

Periodic Mission Review	A NAC directed task to SACEUR to provide the strategic operations assessment of progress towards achieving strategic objectives and their contribution to achievement of the desired end state.
Pilot survey or poll	A scientific tool of study that allows a preliminary analysis before committing to a full blown poll or survey. This is done in order to check the feasibility or to improve the design of the poll or survey.
Predicted Rate of Change (PROC)	The anticipated Rate of Change, which is a useful estimate based on previous history, current intentions and a number of assumptions specific to the circumstances involved.
Priming effect	A previous experience aids the performance of a task without conscious awareness of the previous experience. Exposure to one stimulus (i.e. a question in a survey) may influence a response to another stimulus.
Qualitative	An observation that is a word, or a sentence, or a description, or a code that represents a category (attempting to understand rather than prove).
Quantitative	A number that represents an amount or a count.
Random Sampling	A subset selected from a given population that are chosen randomly, and entirely by chance.
Rate of Change	The amount of change in a metric over a specific time during an operation.
Risk Assessment	The continuous monitoring of strategic and operational risks at the corresponding level of command.
Sample	A subset selected from a given population.
Sampling error	An error that occurs when the statistical characteristics of a population are estimated from a subset of that population.
Subjective	Resulting from an individual's personal opinion, experience and judgement.
Survey	A survey describes a population; it counts 'what is out there'.

System Elements	The functionally, physically or behaviourally related group of regularly interacting or interdependent elements, which forms a unified whole. System elements are specific physical, functional or behavioural entities within a system. Examples of system elements are: people, organisations, facilities, forces, information, processes and conditions. (source: COPD v2)
Systems Analysis	Systems Analysis identifies a network of systems and system elements, their relations and interactions, and their evolution in space and time.
Significant trend	A change has actually occurred and is not in fact due to some other factor, such as normal random variation or simple measurement error.
Trend	A consistent trajectory of change over a number of intervals, or between several groups exhibiting the same kind of change.
Triangulation	The process of using more than two data collection methods in order to examine/double check results. Triangulation helps to overcome problems stemming from reliance upon a single logic, method, source, or set of data.
Undesired Effect	Effects that disrupt or jeopardise the achievement of objectives. (source: COPD v2)
Validation	To show that a methodology is suitable for the purpose it is used.
Validity	An argument is said to be valid to the extent that the conclusions drawn from the data do logically flow from them. Validity can be applied to the validity of the measurement (i.e. extent to which data is accurately measured), the population validity (ie. The extent to which a sample is an accurate representation of the population), the validity of the design (i.e. the extent to which comparisons being made are appropriate to establish the arguments which rest on them).